

Tue Jul 1 09:54:47 2003

us-09-899-295-2.rapm

Ulm, J.  
09/1899295 Page 1  
Seq. ID 2 w/insert

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2003, 16:16:45 ; Search time 311 Seconds

(Without alignments)  
744.242 Million cell updates/sec

Title: US-09-899-295-2

Perfect score: 1860

Sequence: 1 MTLESIMACCLSEAKERR.....VFAVKDTILQNLKEYNLV 359

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main:\*

1: /cgn2\_6/ptodata/1/paa/PCTUS\_COMB.pep.\*

2: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep.\*

3: /cgn2\_6/ptodata/1/paa/US07\_COMB.pep.\*

4: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep.\*

5: /cgn2\_6/ptodata/1/paa/US082\_COMB.pep.\*

6: /cgn2\_6/ptodata/1/paa/US083\_COMB.pep.\*

7: /cgn2\_6/ptodata/1/paa/US084\_COMB.pep.\*

8: /cgn2\_6/ptodata/1/paa/US085\_COMB.pep.\*

9: /cgn2\_6/ptodata/1/paa/US086\_COMB.pep.\*

10: /cgn2\_6/ptodata/1/paa/US087\_COMB.pep.\*

11: /cgn2\_6/ptodata/1/paa/US088\_COMB.pep.\*

12: /cgn2\_6/ptodata/1/paa/US089\_COMB.pep.\*

13: /cgn2\_6/ptodata/1/paa/US090\_COMB.pep.\*

14: /cgn2\_6/ptodata/1/paa/US091\_COMB.pep.\*

15: /cgn2\_6/ptodata/1/paa/US092\_COMB.pep.\*

16: /cgn2\_6/ptodata/1/paa/US093\_COMB.pep.\*

17: /cgn2\_6/ptodata/1/paa/US094\_COMB.pep.\*

18: /cgn2\_6/ptodata/1/paa/US095\_COMB.pep.\*

19: /cgn2\_6/ptodata/1/paa/US096\_COMB.pep.\*

20: /cgn2\_6/ptodata/1/paa/US097\_COMB.pep.\*

21: /cgn2\_6/ptodata/1/paa/US098\_COMB.pep.\*

22: /cgn2\_6/ptodata/1/paa/US099\_COMB.pep.\*

23: /cgn2\_6/ptodata/1/paa/US100\_COMB.pep.\*

24: /cgn2\_6/ptodata/1/paa/US101\_COMB.pep.\*

25: /cgn2\_6/ptodata/1/paa/US102\_COMB.pep.\*

26: /cgn2\_6/ptodata/1/paa/US103\_COMB.pep.\*

27: /cgn2\_6/ptodata/1/paa/US60\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1860	100.0	359	21	US-09-791-537-57305
2	1860	100.0	359	22	US-09-899-295-2
3	1857	99.8	359	18	US-09-471-572-8
4	1857	99.8	359	21	US-09-791-537-81282
5	1849	99.4	359	18	US-09-471-572-7
6	1849	99.4	359	21	US-09-791-537-78690

7	1846	99.2	359	1	PCT-US01-32619-1	Sequence 1, Appl1
8	1846	99.2	359	1	PCT-US01-32619-1	Sequence 1, Appl1
9	1846	99.2	359	23	US-09-984-292-1	Sequence 1, Appl1
10	1846	99.2	359	23	US-09-984-292-1	Sequence 1, Appl1
11	1845	99.1	359	18	US-09-791-537-72283	Sequence 7, Appl1
12	1844	99.1	359	21	US-09-471-572-6	Sequence 6, Appl1
13	1844	99.1	359	21	US-09-791-537-95492	Sequence 95492, A
14	1842	99.0	359	1	PCT-US01-32619-2	Sequence 27, Appl1
15	1842	99.0	359	1	PCT-US01-32619-2	Sequence 37, Appl1
16	1842	99.0	359	21	US-09-791-537-114678	Sequence 114678, A
17	1842	99.0	359	23	US-09-984-292-37	Sequence 37, Appl1
18	1842	99.0	359	23	US-09-984-292-37	Sequence 37, Appl1
19	1837	98.8	359	1	PCT-US01-32619-15	Sequence 15, Appl1
20	1837	98.8	359	1	PCT-US01-32619-15	Sequence 15, Appl1
21	1837	98.8	359	23	US-09-984-292-15	Sequence 15, Appl1
22	1837	98.8	359	23	US-09-984-292-15	Sequence 15, Appl1
23	1830	98.4	359	21	US-09-791-537-78674	Sequence 78674, A
24	1830	98.4	359	21	US-09-791-537-78674	Sequence 78674, A
25	1827	98.2	359	20	US-09-679-664-29	Sequence 29, Appl1
26	1827	98.2	1276	20	US-09-679-664-33	Sequence 33, Appl1
27	1827	98.2	1394	20	US-09-679-664-37	Sequence 37, Appl1
28	1827	98.2	1397	20	US-09-679-664-47	Sequence 47, Appl1
29	1827	98.2	1418	20	US-09-679-664-41	Sequence 41, Appl1
30	1827	98.2	1422	20	US-09-679-664-49	Sequence 49, Appl1
31	1827	98.2	1422	20	US-09-679-664-50	Sequence 50, Appl1
32	1823	98.0	1324	20	US-09-679-664-45	Sequence 45, Appl1
33	1822	98.0	359	21	US-09-791-537-120054	Sequence 120054, A
34	1822	98.0	359	1	PCT-US01-32619-9	Sequence 9, Appl1
35	1822	98.0	359	1	PCT-US01-32619-9	Sequence 9, Appl1
36	1822	98.0	359	23	US-09-984-292-9	Sequence 9, Appl1
37	1822	98.0	359	23	US-09-984-292-9	Sequence 9, Appl1
38	1819	97.8	359	1	PCT-US01-32619-2	Sequence 2, Appl1
39	1819	97.8	359	1	PCT-US01-32619-2	Sequence 2, Appl1
40	1819	97.8	359	23	US-09-984-292-2	Sequence 2, Appl1
41	1819	97.8	359	23	US-09-984-292-2	Sequence 2, Appl1
42	1818	97.7	359	1	PCT-US01-32619-21	Sequence 21, Appl1
43	1818	97.7	359	1	PCT-US01-32619-21	Sequence 21, Appl1
44	1818	97.7	359	23	US-09-984-292-21	Sequence 21, Appl1
45	1818	97.7	359	23	US-09-984-292-21	Sequence 21, Appl1

# ALIGNMENTS

RESULT 1

US-09-791-537-57305

Sequence 57305, Application US/09791537

GENERAL INFORMATION:

APPLICANT: Danzer, Joseph

DEBE, Derek

TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME

TITLE OF INVENTION: METHODS OF USE THEREOF

FILE REFERENCE: 261/210

CURRENT APPLICATION NUMBER: US/09/791,537

CURRENT FILING DATE: 2001-02-22

NUMBER OF SEQ ID NOS: 153055

SOFTWARE: PatentIn version 3.0

SEQ ID NO 57305

LENGTH: 359

TYPE: PRT

ORGANISM: Homo sapiens

US-09-791-537-57305

Query Match	100.0%	Score 1860;	DB 21;	Length 359;
Best Local Similarity	100.0%	Pred. No. 1.1e-175;		
Matches 359;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Oy 1 MTLESIMACCLSEAKERRINDEIRHVRKRDARREKLILGTGSGSKSTFIKMR 60

Db 1 MTLESIMACCLSEAKERRINDEIRHVRKRDARREKLILGTGSGSKSTFIKMR 60

Oy 61 IHGSGVSDCKRGRTKLYQNIPTAMQAMIRAMDTLKIPIKYKHNAQLVREVDVK 120

```
Db 61 IHHSGYSDEBKRGFTKLVYQNIPTAMQAMIRAMDITKIPYKYEHNKAHAQVREVDER 120
QY 121 VSAFENPYDAIKSLMNDPGIOECYDRRREYQSDSTKYYLNDLDRVADPAVLPPTQDVL 180
Db 121 VSAFENPYDAIKSLMNDPGIOECYDRRREYQSDSTKYYLNDLDRVADPAVLPPTQDVL 180
QY 181 RVRVPTGTGIIIEPPDLOSIVFRWVDVGQSRERKWHCEENVTSMIFVALSEYDQVLY 240
Db 181 RVRVPTGTGIIIEPPDLOSIVFRWVDVGQSRERKWHCEENVTSMIFVALSEYDQVLY 240
QY 241 ESDNENMEESKALFRTIITYPWFONSSVILFLNKKDLLEKIMYSHLVDPPEYDGPOR 300
Db 241 ESDNENMEESKALFRTIITYPWFONSSVILFLNKKDLLEKIMYSHLVDPPEYDGPOR 300
QY 301 DAOAAREFIKMFVDLNPDSDKIYSHFTCATDTENIRFVAAVKDTIILQNLKEYNLY 359
Db 301 DAOAAREFIKMFVDLNPDSDKIYSHFTCATDTENIRFVAAVKDTIILQNLKEYNLY 359
```

## RESULT 2

```
US-09-899-295-2
; Sequence 2, Application US/09899295
; GENERAL INFORMATION:
; APPLICANT: Aventis Pharma Deutschland GmbH
; TITLE OF INVENTION: Process for identifying modulators of G protein coupled
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AVE D-2000/A033 englisch
; CURRENT APPLICATION NUMBER: US/09/899,295
; CURRENT FILING DATE: 2001-07-06
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-899-295-2
```

Query Match 100.0%; Score 1860; DB 22; Length 359;

Best Local Similarity 100.0%; Pred. No. 1.1e-175; Matches 359; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MTEESIMACCLSEAKRARRINDEIERHVRDRKRDARELKLILGTGSGSKSTFIKQMR 60
Db 1 MTEESIMACCLSEAKRARRINDEIERHVRDRKRDARELKLILGTGSGSKSTFIKQMR 60
QY 61 IHHSGYSDEBKRGFTKLVYQNIPTAMQAMIRAMDITKIPYKYEHNKAHAQVREVDER 120
Db 61 IHHSGYSDEBKRGFTKLVYQNIPTAMQAMIRAMDITKIPYKYEHNKAHAQVREVDER 120
QY 121 VSAFENPYDAIKSLMNDPGIOECYDRRREYQSDSTKYYLNDLDRVADPAVLPPTQDVL 180
Db 121 VSAFENPYDAIKSLMNDPGIOECYDRRREYQSDSTKYYLNDLDRVADPAVLPPTQDVL 180
QY 181 RVRVPTGTGIIIEPPDLOSIVFRWVDVGQSRERKWHCEENVTSMIFVALSEYDQVLY 240
Db 181 RVRVPTGTGIIIEPPDLOSIVFRWVDVGQSRERKWHCEENVTSMIFVALSEYDQVLY 240
QY 241 ESDNENMEESKALFRTIITYPWFONSSVILFLNKKDLLEKIMYSHLVDPPEYDGPOR 300
Db 241 ESDNENMEESKALFRTIITYPWFONSSVILFLNKKDLLEKIMYSHLVDPPEYDGPOR 300
QY 301 DAOAAREFIKMFVDLNPDSDKIYSHFTCATDTENIRFVAAVKDTIILQNLKEYNLY 359
Db 301 DAOAAREFIKMFVDLNPDSDKIYSHFTCATDTENIRFVAAVKDTIILQNLKEYNLY 359
```

## RESULT 3

```
US-09-471-572-8
; Sequence 8, Application US/09471572
; GENERAL INFORMATION:
; APPLICANT: Jones, Kenneth A.
; APPLICANT: Walker, Mary W.
```

```
APPLICANT: Tamm, Joseph
; APPLICANT: Branchek, Theresa A.
; APPLICANT: Gerald, Christophe P.G.
; TITLE OF INVENTION: Chimeric G-proteins And Uses Thereof
; FILE REFERENCE: 59896
; CURRENT APPLICATION NUMBER: US/09/471,572
; CURRENT FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-471-572-8
```

Query Match 99.8%; Score 1857; DB 18; Length 359;

Best Local Similarity 99.7%; Pred. No. 2.1e-175; Matches 358; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MTEESIMACCLSEAKRARRINDEIERHVRDRKRDARELKLILGTGSGSKSTFIKQMR 60
Db 1 MTEESIMACCLSEAKRARRINDEIERHVRDRKRDARELKLILGTGSGSKSTFIKQMR 60
QY 61 IHHSGYSDEBKRGFTKLVYQNIPTAMQAMIRAMDITKIPYKYEHNKAHAQVREVDER 120
Db 61 IHHSGYSDEBKRGFTKLVYQNIPTAMQAMIRAMDITKIPYKYEHNKAHAQVREVDER 120
QY 121 VSAFENPYDAIKSLMNDPGIOECYDRRREYQSDSTKYYLNDLDRVADPAVLPPTQDVL 180
Db 121 VSAFENPYDAIKSLMNDPGIOECYDRRREYQSDSTKYYLNDLDRVADPAVLPPTQDVL 180
QY 181 RVRVPTGTGIIIEPPDLOSIVFRWVDVGQSRERKWHCEENVTSMIFVALSEYDQVLY 240
Db 181 RVRVPTGTGIIIEPPDLOSIVFRWVDVGQSRERKWHCEENVTSMIFVALSEYDQVLY 240
QY 241 ESDNENMEESKALFRTIITYPWFONSSVILFLNKKDLLEKIMYSHLVDPPEYDGPOR 300
Db 241 ESDNENMEESKALFRTIITYPWFONSSVILFLNKKDLLEKIMYSHLVDPPEYDGPOR 300
QY 301 DAOAAREFIKMFVDLNPDSDKIYSHFTCATDTENIRFVAAVKDTIILQNLKEYNLY 359
Db 301 DAOAAREFIKMFVDLNPDSDKIYSHFTCATDTENIRFVAAVKDTIILQNLKEYNLY 359
```

## RESULT 4

```
US-09-791-537-81282
; Sequence 81282, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 81282
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-791-537-81282
```

Query Match 99.8%; Score 1857; DB 21; Length 359;

Best Local Similarity 99.7%; Pred. No. 2.1e-175; Matches 358; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MTEESIMACCLSEAKRARRINDEIERHVRDRKRDARELKLILGTGSGSKSTFIKQMR 60
Db 1 MTEESIMACCLSEAKRARRINDEIERHVRDRKRDARELKLILGTGSGSKSTFIKQMR 60
QY 61 IHHSGYSDEBKRGFTKLVYQNIPTAMQAMIRAMDITKIPYKYEHNKAHAQVREVDER 120
```

```
Db 61 I H G S G Y S D E D K R G F T K L V Y Q N I F T A M Q A M I R A M D T L K I P Y K Y E H N K A H Q L V R E Y D V E K 120
Qy 121 V S A F E N P Y V D A I S L M N D P G I O E C Y D R R R E Y Q L S D S T K Y Y L N D L D R A D A P A Y L P T O O D V L 180
Db 121 V S A F E N P Y V D A I S L M N D P G I O E C Y D R R R E Y Q L S D S T K Y Y L N D L D R A D A P A Y L P T O O D V L 180
Qy 181 R V A P P T G I I E Y P F D O S V I F R A W D V G G O S E R R K W I H C E N T S I M F V A L S E Y D O V L V 240
Db 181 R V A P P T G I I E Y P F D O S V I F R A W D V G G O S E R R K W I H C E N T S I M F V A L S E Y D O V L V 240
Qy 241 E S D N E N M E S K A L F R I I T Y P W F O N S V I L F L N K K D L E E K I M Y S H L V D Y P P E Y D G P O R 300
Db 241 E S D N E N M E S K A L F R I I T Y P W F O N S V I L F L N K K D L E E K I M Y S H L V D Y P P E Y D G P O R 300
Qy 301 D A Q A A R E F I L K M E V D L N P D S D K I I Y S H F T C A T D T E N I R F V A A V K D T I I O L N L K E Y N L V 359
Db 301 D A Q A A R E F I L K M E V D L N P D S D K I I Y S H F T C A T D T E N I R F V A A V K D T I I O L N L K E Y N L V 359

RESULT 5
US-09-471-572-7
; Sequence 7, Application US/09471572
; GENERAL INFORMATION:
; APPLICANT: Jones, Kenneth A.
; APPLICANT: Walker, Mary W.
; APPLICANT: Tamm, Joseph
; APPLICANT: Branchek, Theresa A.
; APPLICANT: Gerald, Christophe P.G.
; TITLE OF INVENTION: Chimeric G-Proteins And Uses Thereof
; FILE REFERENCE: 59896
; CURRENT APPLICATION NUMBER: US/09/471, 572
; CURRENT FILING DATE: 1999-12-23
; NUMBER OF SEQ. ID NOS: 45
; SOFTWARE: Patentln Ver. 2.1
; SEQ. ID NO. 7
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Canis familiaris
US-09-471-572-7

Query Match 99.4%; Score 1849; DB 18; Length 359;
Best Local Similarity 99.4%; Pred. No. 1.3e-174;
Matches 357; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 M T L E S I M A C C L S E A K E A R I N D E I E R H V R D K R D A R R E L K L L L G T G S G K S T F T K O M R 60
Db 1 M T L E S I M A C C L S E A K E A R I N D E I E R Q L R D K R D A R R E L K L L L G T G S G K S T F T K O M R 60
Qy 61 I I H G S G Y S D E D K R G F T K L V Y Q N I F T A M Q A M I R A M D T L K I P Y K Y E H N K A H Q L V R E Y D V E K 120
Db 61 I I H G S G Y S D E D K R G F T K L V Y Q N I F T A M Q A M I R A M D T L K I P Y K Y E H N K A H Q L V R E Y D V E K 120
Qy 121 V S A F E N P Y V D A I S L M N D P G I O E C Y D R R R E Y Q L S D S T K Y Y L N D L D R A D A P A Y L P T O O D V L 180
Db 121 V S A F E N P Y V D A I S L M N D P G I O E C Y D R R R E Y Q L S D S T K Y Y L N D L D R A D A P A Y L P T O O D V L 180
Qy 181 R V A P P T G I I E Y P F D O S V I F R A W D V G G O S E R R K W I H C E N T S I M F V A L S E Y D O V L V 240
Db 181 R V A P P T G I I E Y P F D O S V I F R A W D V G G O S E R R K W I H C E N T S I M F V A L S E Y D O V L V 240
Qy 241 E S D N E N M E S K A L F R I I T Y P W F O N S V I L F L N K K D L E E K I M Y S H L V D Y P P E Y D G P O R 300
Db 241 E S D N E N M E S K A L F R I I T Y P W F O N S V I L F L N K K D L E E K I M Y S H L V D Y P P E Y D G P O R 300
Qy 301 D A Q A A R E F I L K M E V D L N P D S D K I I Y S H F T C A T D T E N I R F V A A V K D T I I O L N L K E Y N L V 359
Db 301 D A Q A A R E F I L K M E V D L N P D S D K I I Y S H F T C A T D T E N I R F V A A V K D T I I O L N L K E Y N L V 359

RESULT 6
US-09-791-537-78690
; Sequence 78690, Application US/09791537
```

```
; GENERAL INFORMATION:
; APPLICANT: Biomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ. ID NOS: 153055
; SOFTWARE: Patentln version 3.0
; SEQ. ID NO 78690
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-791-537-78690

Query Match 99.4%; Score 1849; DB 21; Length 359;
Best Local Similarity 99.4%; Pred. No. 1.3e-174;
Matches 357; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 M T L E S I M A C C L S E A K E A R I N D E I E R H V R D K R D A R R E L K L L L G T G S G K S T F T K O M R 60
Db 1 M T L E S I M A C C L S E A K E A R I N D E I E R Q L R D K R D A R R E L K L L L G T G S G K S T F T K O M R 60
Qy 61 I I H G S G Y S D E D K R G F T K L V Y Q N I F T A M Q A M I R A M D T L K I P Y K Y E H N K A H Q L V R E Y D V E K 120
Db 61 I I H G S G Y S D E D K R G F T K L V Y Q N I F T A M Q A M I R A M D T L K I P Y K Y E H N K A H Q L V R E Y D V E K 120
Qy 121 V S A F E N P Y V D A I S L M N D P G I O E C Y D R R R E Y Q L S D S T K Y Y L N D L D R A D A P A Y L P T O O D V L 180
Db 121 V S A F E N P Y V D A I S L M N D P G I O E C Y D R R R E Y Q L S D S T K Y Y L N D L D R A D A P A Y L P T O O D V L 180
Qy 181 R V A P P T G I I E Y P F D O S V I F R A W D V G G O S E R R K W I H C E N T S I M F V A L S E Y D O V L V 240
Db 181 R V A P P T G I I E Y P F D O S V I F R A W D V G G O S E R R K W I H C E N T S I M F V A L S E Y D O V L V 240
Qy 241 E S D N E N M E S K A L F R I I T Y P W F O N S V I L F L N K K D L E E K I M Y S H L V D Y P P E Y D G P O R 300
Db 241 E S D N E N M E S K A L F R I I T Y P W F O N S V I L F L N K K D L E E K I M Y S H L V D Y P P E Y D G P O R 300
Qy 301 D A Q A A R E F I L K M E V D L N P D S D K I I Y S H F T C A T D T E N I R F V A A V K D T I I O L N L K E Y N L V 359
Db 301 D A Q A A R E F I L K M E V D L N P D S D K I I Y S H F T C A T D T E N I R F V A A V K D T I I O L N L K E Y N L V 359

RESULT 7
PCT-US01-32619-1
; Sequence 1, Application PC/TUS0132619
; GENERAL INFORMATION:
; APPLICANT: SENOMYX, INC.
; TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE
; TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF
; TITLE OF INVENTION: CHEMOSENSORY RECEPTORS
; FILE REFERENCE: 078003-140309
; CURRENT APPLICATION NUMBER: PCT/US01/32619
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: 60/243,770
; PRIOR FILING DATE: 2000-10-30
; NUMBER OF SEQ. ID NOS: 31
; SOFTWARE: Patentln Ver. 2.1
; SEQ. ID NO 1
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Mus sp.
PCT-US01-32619-1

Query Match 99.2%; Score 1846; DB 1; Length 359;
Best Local Similarity 99.2%; Pred. No. 2.6e-174;
Matches 356; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

Db 1 MTELSIMACCLSEAKRINDEIREROLRDRKDRARREKLILLGTGSGKSTFTKQMR 60  
QY 61 IHGSGYSDKRGKFTLYQNIPTAMQAMIRAMDTLKIPYKEHNKAHQVREYDVK 120  
Db 61 IHGSGYSDKRGKFTLYQNIPTAMQAMIRAMDTLKIPYKEHNKAHQVREYDVK 120  
QY 121 VSAFENPYDAIKSLNNDPGIOECYDRRREYQSDSTKYLYNDLDRVADPAYLPTQOVL 180  
Db 121 VSAFENPYDAIKSLNNDPGIOECYDRRREYQSDSTKYLYNDLDRVADPAYLPTQOVL 180  
QY 181 RVAVPTGTIIEYFPDQSVIFRWVDVGGORSEKRNHCEFNVTSLMFLVASEYDOVL 240  
Db 181 RVAVPTGTIIEYFPDQSVIFRWVDVGGORSEKRNHCEFNVTSLMFLVASEYDOVL 240  
QY 241 ESDNENRMEESKALFTITITYPWFONSSVILFLNKDLEEKIMYSHLVDPPEYDGPOR 300  
Db 241 ESDNENRMEESKALFTITITYPWFONSSVILFLNKDLEEKIMYSHLVDPPEYDGPOR 300  
QY 301 DAQAREFLIKMFVDLNPDSKTIYSHFTCATDTENIRVFAAVKDTIILQNLKEYNLV 359  
Db 301 DAQAREFLIKMFVDLNPDSKTIYSHFTCATDTENIRVFAAVKDTIILQNLKEYNLV 359

## RESULT 8

PCT-US01-32819-1  
; Sequence 1, Application PC/YUS0132819  
; GENERAL INFORMATION:  
; APPLICANT: YAO, YONG  
; APPLICANT: XU, HONG  
; TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE  
; TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF  
; TITLE OF INVENTION: CHEMOSENSORY RECEPTORS  
; FILE REFERENCE: 078003-0280649  
; CURRENT APPLICATION NUMBER: PCT/US01/32819  
; PRIOR FILING DATE: 2002-03-07  
; PRIOR APPLICATION NUMBER: 60/243,770  
; PRIOR FILING DATE: 2000-10-30  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 359  
; TYPE: PRT  
; ORGANISM: Mus sp.  
PCT-US01-32819-1

Query Match 99.2%; Score 1846; DB 1; Length 359;  
Best Local Similarity 99.2%; Pred. No. 2.6e-174;  
Matches 356; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MTELSIMACCLSEAKRINDEIREROLRDRKDRARREKLILLGTGSGKSTFTKQMR 60  
Db 1 MTELSIMACCLSEAKRINDEIREROLRDRKDRARREKLILLGTGSGKSTFTKQMR 60  
QY 61 IHGSGYSDKRGKFTLYQNIPTAMQAMIRAMDTLKIPYKEHNKAHQVREYDVK 120  
Db 61 IHGSGYSDKRGKFTLYQNIPTAMQAMIRAMDTLKIPYKEHNKAHQVREYDVK 120  
QY 121 VSAFENPYDAIKSLNNDPGIOECYDRRREYQSDSTKYLYNDLDRVADPAYLPTQOVL 180  
Db 121 VSAFENPYDAIKSLNNDPGIOECYDRRREYQSDSTKYLYNDLDRVADPAYLPTQOVL 180  
QY 181 RVAVPTGTIIEYFPDQSVIFRWVDVGGORSEKRNHCEFNVTSLMFLVASEYDOVL 240  
Db 181 RVAVPTGTIIEYFPDQSVIFRWVDVGGORSEKRNHCEFNVTSLMFLVASEYDOVL 240  
QY 241 ESDNENRMEESKALFTITITYPWFONSSVILFLNKDLEEKIMYSHLVDPPEYDGPOR 300  
Db 241 ESDNENRMEESKALFTITITYPWFONSSVILFLNKDLEEKIMYSHLVDPPEYDGPOR 300  
QY 301 DAQAREFLIKMFVDLNPDSKTIYSHFTCATDTENIRVFAAVKDTIILQNLKEYNLV 359  
Db 301 DAQAREFLIKMFVDLNPDSKTIYSHFTCATDTENIRVFAAVKDTIILQNLKEYNLV 359

## RESULT 9

US-09-984-292-1  
; Sequence 1, Application US/09984292  
; GENERAL INFORMATION:  
; APPLICANT: YAO, YONG  
; APPLICANT: XU, HONG  
; TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE  
; TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF  
; FILE REFERENCE: 078003-0280649  
; CURRENT APPLICATION NUMBER: US/09/984,292  
; PRIOR FILING DATE: 2001-10-29  
; PRIOR APPLICATION NUMBER: 60/243,770  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 359  
; TYPE: PRT  
; ORGANISM: Mus sp.  
US-09-984-292-1

Query Match 99.2%; Score 1846; DB 23; Length 359;  
Best Local Similarity 99.2%; Pred. No. 2.6e-174;  
Matches 356; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MTELSIMACCLSEAKRINDEIREROLRDRKDRARREKLILLGTGSGKSTFTKQMR 60  
Db 1 MTELSIMACCLSEAKRINDEIREROLRDRKDRARREKLILLGTGSGKSTFTKQMR 60  
QY 61 IHGSGYSDKRGKFTLYQNIPTAMQAMIRAMDTLKIPYKEHNKAHQVREYDVK 120  
Db 61 IHGSGYSDKRGKFTLYQNIPTAMQAMIRAMDTLKIPYKEHNKAHQVREYDVK 120  
QY 121 VSAFENPYDAIKSLNNDPGIOECYDRRREYQSDSTKYLYNDLDRVADPAYLPTQOVL 180  
Db 121 VSAFENPYDAIKSLNNDPGIOECYDRRREYQSDSTKYLYNDLDRVADPAYLPTQOVL 180  
QY 181 RVAVPTGTIIEYFPDQSVIFRWVDVGGORSEKRNHCEFNVTSLMFLVASEYDOVL 240  
Db 181 RVAVPTGTIIEYFPDQSVIFRWVDVGGORSEKRNHCEFNVTSLMFLVASEYDOVL 240  
QY 241 ESDNENRMEESKALFTITITYPWFONSSVILFLNKDLEEKIMYSHLVDPPEYDGPOR 300  
Db 241 ESDNENRMEESKALFTITITYPWFONSSVILFLNKDLEEKIMYSHLVDPPEYDGPOR 300  
QY 301 DAQAREFLIKMFVDLNPDSKTIYSHFTCATDTENIRVFAAVKDTIILQNLKEYNLV 359  
Db 301 DAQAREFLIKMFVDLNPDSKTIYSHFTCATDTENIRVFAAVKDTIILQNLKEYNLV 359

## RESULT 10

US-09-989-497-1  
; Sequence 1, Application US/09989497  
; GENERAL INFORMATION:  
; APPLICANT: YAO, YONG  
; APPLICANT: XU, HONG  
; TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE  
; TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF  
; FILE REFERENCE: 078003-0280735  
; CURRENT APPLICATION NUMBER: US/09/989,497  
; PRIOR FILING DATE: 2001-11-21  
; PRIOR APPLICATION NUMBER: 09/984,292  
; PRIOR FILING DATE: 2001-10-29  
; PRIOR APPLICATION NUMBER: 60/243,770  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 359  
; TYPE: PRT

ORGANISM: Mus sp.  
US-09-989-497-1

Query Match 99.2%; Score 1846; DB 23; Length 359;  
Best Local Similarity 99.2%; Pred. No. 2.6e-174;  
Matches 356; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLESIMACCLSEEAKEARRINDEIERHVRDRDARRELKLLLGTEGSGSTFIKQMR 60  
DB 1 MLESIMACCLSEEAKEARRINDEIERHVRDRDARRELKLLLGTEGSGSTFIKQMR 60  
QY 61 IIHGSGYDEDEKRGFTKLVYONIFTAQAMIRAMDTLKIPYKEHNKAHAQLVREVDYK 120  
DB 61 IIHGSGYDEDEKRGFTKLVYONIFTAQAMIRAMDTLKIPYKEHNKAHAQLVREVDYK 120  
QY 121 VSAFENPYVDAIKSLMNDPGIOECYDRREYQLSDSTKYLLNDLDRVADPALPTQOQVYL 180  
DB 121 VSAFENPYVDAIKSLMNDPGIOECYDRREYQLSDSTKYLLNDLDRVADPALPTQOQVYL 180  
QY 181 RVRVPTGTGIIIEYFPDQSVIFRMVDVGGORSEERKMHICFENVTSIMFLVALSEYDQVYL 240  
DB 181 RVRVPTGTGIIIEYFPDQSVIFRMVDVGGORSEERKMHICFENVTSIMFLVALSEYDQVYL 240  
QY 241 ESDNENRMEESKALFRITITTPWFONSSVILFLNKKDLLEEKIMVSHLVDFPEYDGPQR 300  
DB 241 ESDNENRMEESKALFRITITTPWFONSSVILFLNKKDLLEEKIMVSHLVDFPEYDGPQR 300  
QY 301 DAQAAREFLKMFVLDNPDSDKIISHTFCATDENTINIRVFAAVDITLQNLKEYNLV 359  
DB 301 DAQAAREFLKMFVLDNPDSDKIISHTFCATDENTINIRVFAAVDITLQNLKEYNLV 359

## RESULT 11

US-09-791-537-72283  
; Sequence 72283, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Biomomix, Inc.  
; APPLICANT: Debe, Derek  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 72283  
; LENGTH: 359  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-09-791-537-72283

Query Match 99.2%; Score 1845; DB 21; Length 359;  
Best Local Similarity 98.9%; Pred. No. 3.3e-174;  
Matches 355; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLESIMACCLSEEAKEARRINDEIERHVRDRDARRELKLLLGTEGSGSTFIKQMR 60  
DB 1 MLESIMACCLSEEAKEARRINDEIERHVRDRDARRELKLLLGTEGSGSTFIKQMR 60  
QY 61 IIHGSGYDEDEKRGFTKLVYONIFTAQAMIRAMDTLKIPYKEHNKAHAQLVREVDYK 120  
DB 61 IIHGSGYDEDEKRGFTKLVYONIFTAQAMIRAMDTLKIPYKEHNKAHAQLVREVDYK 120  
QY 121 VSAFENPYVDAIKSLMNDPGIOECYDRREYQLSDSTKYLLNDLDRVADPALPTQOQVYL 180  
DB 121 VSAFENPYVDAIKSLMNDPGIOECYDRREYQLSDSTKYLLNDLDRVADPALPTQOQVYL 180  
QY 181 RVRVPTGTGIIIEYFPDQSVIFRMVDVGGORSEERKMHICFENVTSIMFLVALSEYDQVYL 240  
DB 181 RVRVPTGTGIIIEYFPDQSVIFRMVDVGGORSEERKMHICFENVTSIMFLVALSEYDQVYL 240  
QY 241 ESDNENRMEESKALFRITITTPWFONSSVILFLNKKDLLEEKIMVSHLVDFPEYDGPQR 300

DB 241 ESDNENRMEESKALFRITITTPWFONSSVILFLNKKDLLEEKIMVSHLVDFPEYDGPQR 300  
QY 301 DAQAAREFLKMFVLDNPDSDKIISHTFCATDENTINIRVFAAVDITLQNLKEYNLV 359  
DB 301 DAQAAREFLKMFVLDNPDSDKIISHTFCATDENTINIRVFAAVDITLQNLKEYNLV 359

## RESULT 12

US-09-471-572-6  
; Sequence 6, Application US/09471572  
; GENERAL INFORMATION:  
; APPLICANT: Jones, Kenneth A.  
; APPLICANT: Walker, Mary W.  
; APPLICANT: Tamm, Joseph  
; APPLICANT: Branchek, Theresa A.  
; TITLE OF INVENTION: Chimeric G-Proteins And Uses Thereof  
; FILE REFERENCE: 59896  
; CURRENT APPLICATION NUMBER: US/09/471,572  
; CURRENT FILING DATE: 1999-12-23  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 359  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-471-572-6

Query Match 99.1%; Score 1844; DB 18; Length 359;  
Best Local Similarity 99.2%; Pred. No. 4.1e-174;  
Matches 356; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MLESIMACCLSEEAKEARRINDEIERHVRDRDARRELKLLLGTEGSGSTFIKQMR 60  
DB 1 MLESIMACCLSEEAKEARRINDEIERHVRDRDARRELKLLLGTEGSGSTFIKQMR 60  
QY 61 IIHGSGYDEDEKRGFTKLVYONIFTAQAMIRAMDTLKIPYKEHNKAHAQLVREVDYK 120  
DB 61 IIHGSGYDEDEKRGFTKLVYONIFTAQAMIRAMDTLKIPYKEHNKAHAQLVREVDYK 120  
QY 121 VSAFENPYVDAIKSLMNDPGIOECYDRREYQLSDSTKYLLNDLDRVADPALPTQOQVYL 180  
DB 121 VSAFENPYVDAIKSLMNDPGIOECYDRREYQLSDSTKYLLNDLDRVADPALPTQOQVYL 180  
QY 181 RVRVPTGTGIIIEYFPDQSVIFRMVDVGGORSEERKMHICFENVTSIMFLVALSEYDQVYL 240  
DB 181 RVRVPTGTGIIIEYFPDQSVIFRMVDVGGORSEERKMHICFENVTSIMFLVALSEYDQVYL 240  
QY 241 ESDNENRMEESKALFRITITTPWFONSSVILFLNKKDLLEEKIMVSHLVDFPEYDGPQR 300  
DB 241 ESDNENRMEESKALFRITITTPWFONSSVILFLNKKDLLEEKIMVSHLVDFPEYDGPQR 300  
QY 301 DAQAAREFLKMFVLDNPDSDKIISHTFCATDENTINIRVFAAVDITLQNLKEYNLV 359  
DB 301 DAQAAREFLKMFVLDNPDSDKIISHTFCATDENTINIRVFAAVDITLQNLKEYNLV 359

## RESULT 13

US-09-791-537-95492  
; Sequence 95492, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Biomomix, Inc.  
; APPLICANT: Debe, Derek  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 95492



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: June 30, 2003, 16:18:06 ; Search time 50 Seconds  
(without alignments)  
1870.039 Million cell updates/sec

Title: US-09-899-295-2

Perfect score: 1860

Sequence: 1 MTEESIMACCLSEAKREARR.....VFAVKPTIQLNKEYNLV 359

## Scoring table:

BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1197705 seqs, 260451061 residues

Total number of hits satisfying chosen parameters: 1197705

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

## Database :

Pending\_Patents\_AA\_New:\*  
1: /cgn2\_6/ptodata/2/paa/PCR\_NEW\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep:\*  
7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	1849	99.4	359	6	US-10-405-027-3227	Sequence 3227, Ap
2	1849	99.4	359	7	US-60-453-135-10419	Sequence 10419, A
3	1849	99.4	359	7	US-60-453-050-10419	Sequence 10419, A
4	1849	99.4	359	7	US-60-455-444-5597	Sequence 5597, Ap
5	1849	99.4	359	7	US-60-465-241-5597	Sequence 5597, Ap
6	1849	99.4	359	7	US-60-466-412-10419	Sequence 10419, A
7	1844	99.1	359	6	US-10-238-361-2	Sequence 2, Appl
8	1817	97.7	353	6	PCT-US03-02452-13	Sequence 13, Appl
9	1817	97.7	353	6	US-10-352-843-13	Sequence 13, Appl
10	1687	90.7	387	5	US-09-724-676-49632	Sequence 49632, A
11	1687	90.7	387	5	US-09-724-676-49632	Sequence 49632, A
12	1657	89.1	351	5	US-09-949-016-7853	Sequence 7853, Ap
13	1535	82.5	355	6	US-10-326-896-2	Sequence 2, Appl
14	1513	81.3	355	6	PCT-US03-02452-12	Sequence 12, Appl
15	1513	81.3	355	6	US-10-369-493-4982	Sequence 4982, Ap
16	1513	81.3	355	6	US-10-352-843-12	Sequence 12, Appl
17	1104	59.4	252	5	US-09-949-016-9408	Sequence 9408, Ap
18	1066	57.3	374	5	US-09-468-002-4	Sequence 4, Appl
19	1066	57.3	374	6	US-10-319-416-2	Sequence 2, Appl
20	1055	56.7	374	6	US-10-299-442-20	Sequence 20, Appl
21	1055	56.7	374	6	US-10-414-797-20	Sequence 20, Appl
22	1044	56.1	374	5	US-09-468-002-2	Sequence 2, Appl
23	1044	56.1	378	5	US-09-949-016-7851	Sequence 7851, Ap
24	1033	55.5	374	6	US-10-319-416-4	Sequence 4, Appl
25	898.5	48.3	354	1	PCT-US03-02452-14	Sequence 14, Appl
26	898.5	48.3	354	6	US-10-352-843-14	Sequence 14, Appl

27	897.5	48.3	354	6	US-10-180-930-2	Sequence 2, Appl
28	893.5	48.0	353	6	US-10-219-051B-3379	Sequence 3379, Ap
29	893.5	48.0	353	6	US-10-219-051B-3383	Sequence 3383, Ap
30	892.5	48.0	353	6	US-10-219-051B-3377	Sequence 3377, Ap
31	892.5	48.0	353	6	US-10-219-051B-3381	Sequence 3381, Ap
32	888	47.7	710	1	PCT-US02-31059B-12	Sequence 12, Appl
33	881.5	47.4	354	5	US-09-949-016-6727	Sequence 6727, Ap
34	881.5	47.4	354	7	US-60-443-566-3898	Sequence 3898, Ap
35	881.5	47.4	354	7	US-60-452-680-23965	Sequence 23965, A
36	881.5	47.4	354	7	US-60-453-135-14917	Sequence 14917, A
37	881.5	47.4	354	7	US-60-453-050-14917	Sequence 14917, A
38	881.5	47.4	354	7	US-60-455-444-8098	Sequence 8098, Ap
39	881.5	47.4	354	7	US-60-465-241-8098	Sequence 8098, Ap
40	881.5	47.4	354	7	US-60-466-412-14917	Sequence 14917, A
41	877.5	47.2	353	6	US-10-219-051B-11687	Sequence 11687, A
42	877	47.2	355	7	US-60-453-135-9379	Sequence 9379, Ap
43	877	47.2	355	7	US-60-453-050-9379	Sequence 9379, Ap
44	877	47.2	355	7	US-60-466-412-9379	Sequence 9379, Ap
45	877	47.2	395	5	US-09-949-016-11560	Sequence 11560, A

## ALIGNMENTS

RESULT 1  
US-10-405-027-3227  
Sequence 3227, Application US/10405027  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Human Secreted Proteins  
FILE REFERENCE: PS806P1  
CURRENT APPLICATION NUMBER: US/10/405,027  
CURRENT FILING DATE: 2003-04-07  
PRIOR APPLICATION NUMBER: 60/369,608  
PRIOR FILING DATE: 2002-04-04  
PRIOR APPLICATION NUMBER: 60/376,175  
PRIOR FILING DATE: 2002-04-30  
NUMBER OF SEQ ID NOS: 5810  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 3227  
LENGTH: 359  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-405-027-3227

Query Match	99.4%	Score 1849:	DB 6:	Length 359:
Best Local Similarity	99.4%	Pred. NO. 4.1e-183;		
Matches 357:	Conservative 1:	Mismatches 1:	Indels 0:	Gaps 0:
OY	1	MTEESIMACCLSEAKREARRINDEIERHVRDKDARRELKLLLGTEGSGSTFIKQMR	60	
DB	1	MTEESIMACCLSEAKREARRINDEIERHVRDKDARRELKLLLGTEGSGSTFIKQMR	60	
OY	61	IIHSGSYDEDKRGFTKLVYONITFAMQAMTRAMDYKIPYKHEHNAHQAQVREVDYK	120	
DB	61	IIHSGSYDEDKRGFTKLVYONITFAMQAMTRAMDYKIPYKHEHNAHQAQVREVDYK	120	
OY	121	VSAFENPVDAIKSLMNDPGIOECYDRREYQSLDSITVYVYNDIDRVADPVPYLPQOVL	180	
DB	121	VSAFENPVDAIKSLMNDPGIOECYDRREYQSLDSITVYVYNDIDRVADPVPYLPQOVL	180	
OY	181	RVRPPTGILYPPDLQSVIRFMDVGCORREKRWIHCFENVTSIMFLVALSEYDVLV	240	
DB	181	RVRPPTGILYPPDLQSVIRFMDVGCORREKRWIHCFENVTSIMFLVALSEYDVLV	240	
OY	241	ESDENRMEESKALFRITITVYVWFONSSVILFENKDLLEKIMYSHLVDFPEYDQOR	300	
DB	241	ESDENRMEESKALFRITITVYVWFONSSVILFENKDLLEKIMYSHLVDFPEYDQOR	300	
OY	301	DAQAREFILLMFVDLNDSDSKITVYSHETCATDENTREVFNAVKDTIQLNKEYNLV	359	
DB	301	DAQAREFILLMFVDLNDSDSKITVYSHETCATDENTREVFNAVKDTIQLNKEYNLV	359	

RESULT 2  
US-60-453-135-10419  
Sequence 10419, Application US/60453135  
GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele  
APPLICANT: IAKOUBOVA, Olga  
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
FILE REFERENCE: C1001456  
CURRENT APPLICATION NUMBER: US/60/453,135  
CURRENT FILING DATE: 2003-03-10  
NUMBER OF SEQ ID NOS: 82762  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 10419  
LENGTH: 359  
TYPE: PRF  
ORGANISM: Homo sapiens  
US-60-453-135-10419

Query Match 99.4%; Score 1849; DB 7; Length 359;  
Best Local Similarity 99.4%; Pred. No. 4.1e-183;  
Matches 357; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MTEESIMACCLSEAKARRINDEIERHVRDRKRDARRELKLLLGTEGSGKSTFIKQMR 60  
DB 1 MTEESIMACCLSEAKARRINDEIERHVRDRKRDARRELKLLLGTEGSGKSTFIKQMR 60  
QY 61 IHHSGYSDDEKRGFTKLVYQNIPTAMQAMIRAMDTLKIPYKXENKHAQOLVREYDVK 120  
DB 61 IHHSGYSDDEKRGFTKLVYQNIPTAMQAMIRAMDTLKIPYKXENKHAQOLVREYDVK 120  
QY 121 VSAFENPYVDAIKSLMNDPGIOECYDRREYQSDSTKYLLNDLDRADPAVLPYDQDVL 180  
DB 121 VSAFENPYVDAIKSLMNDPGIOECYDRREYQSDSTKYLLNDLDRADPAVLPYDQDVL 180  
QY 181 RVAVPTGIIIEYFPDLOSIFRWDVGGQSRERKWHCFENVTSMFLVALSEYDQVYL 240  
DB 181 RVAVPTGIIIEYFPDLOSIFRWDVGGQSRERKWHCFENVTSMFLVALSEYDQVYL 240  
QY 241 ESDNENRMEESKALFRTIITYPWFONSSVILFLNKKDLLEKIMYSHLVDFPEYDGPQR 300  
DB 241 ESDNENRMEESKALFRTIITYPWFONSSVILFLNKKDLLEKIMYSHLVDFPEYDGPQR 300  
QY 301 DAQAAREFLIKMFVDLNPDSKIIYSHFTCATDTENIRFVAAYKDTIILQNLKEYNLY 359  
DB 301 DAQAAREFLIKMFVDLNPDSKIIYSHFTCATDTENIRFVAAYKDTIILQNLKEYNLY 359

RESULT 3  
US-60-453-050-10419  
Sequence 10419, Application US/60453050  
GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele  
APPLICANT: LUKE, May  
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
FILE REFERENCE: C1001457  
CURRENT APPLICATION NUMBER: US/60/453,050  
CURRENT FILING DATE: 2003-03-10  
NUMBER OF SEQ ID NOS: 82762  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 10419  
LENGTH: 359  
TYPE: PRF  
ORGANISM: Homo sapiens  
US-60-453-050-10419

Query Match 99.4%; Score 1849; DB 7; Length 359;  
Best Local Similarity 99.4%; Pred. No. 4.1e-183;  
Matches 357; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MTEESIMACCLSEAKARRINDEIERHVRDRKRDARRELKLLLGTEGSGKSTFIKQMR 60

DB 1 MTEESIMACCLSEAKARRINDEIERHVRDRKRDARRELKLLLGTEGSGKSTFIKQMR 60  
QY 61 IHHSGYSDDEKRGFTKLVYQNIPTAMQAMIRAMDTLKIPYKXENKHAQOLVREYDVK 120  
DB 61 IHHSGYSDDEKRGFTKLVYQNIPTAMQAMIRAMDTLKIPYKXENKHAQOLVREYDVK 120  
QY 121 VSAFENPYVDAIKSLMNDPGIOECYDRREYQSDSTKYLLNDLDRADPAVLPYDQDVL 180  
DB 121 VSAFENPYVDAIKSLMNDPGIOECYDRREYQSDSTKYLLNDLDRADPAVLPYDQDVL 180  
QY 181 RVAVPTGIIIEYFPDLOSIFRWDVGGQSRERKWHCFENVTSMFLVALSEYDQVYL 240  
DB 181 RVAVPTGIIIEYFPDLOSIFRWDVGGQSRERKWHCFENVTSMFLVALSEYDQVYL 240  
QY 241 ESDNENRMEESKALFRTIITYPWFONSSVILFLNKKDLLEKIMYSHLVDFPEYDGPQR 300  
DB 241 ESDNENRMEESKALFRTIITYPWFONSSVILFLNKKDLLEKIMYSHLVDFPEYDGPQR 300  
QY 301 DAQAAREFLIKMFVDLNPDSKIIYSHFTCATDTENIRFVAAYKDTIILQNLKEYNLY 359  
DB 301 DAQAAREFLIKMFVDLNPDSKIIYSHFTCATDTENIRFVAAYKDTIILQNLKEYNLY 359

RESULT 4  
US-60-455-444-5597  
Sequence 5597, Application US/60455444  
GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele  
APPLICANT: BEGOVICH, Ann  
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
FILE REFERENCE: C1001455  
CURRENT APPLICATION NUMBER: US/60/455,444  
CURRENT FILING DATE: 2003-03-18  
NUMBER OF SEQ ID NOS: 50986  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 5597  
LENGTH: 359  
TYPE: PRF  
ORGANISM: Homo sapiens  
US-60-455-444-5597

Query Match 99.4%; Score 1849; DB 7; Length 359;  
Best Local Similarity 99.4%; Pred. No. 4.1e-183;  
Matches 357; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MTEESIMACCLSEAKARRINDEIERHVRDRKRDARRELKLLLGTEGSGKSTFIKQMR 60  
DB 1 MTEESIMACCLSEAKARRINDEIERHVRDRKRDARRELKLLLGTEGSGKSTFIKQMR 60  
QY 61 IHHSGYSDDEKRGFTKLVYQNIPTAMQAMIRAMDTLKIPYKXENKHAQOLVREYDVK 120  
DB 61 IHHSGYSDDEKRGFTKLVYQNIPTAMQAMIRAMDTLKIPYKXENKHAQOLVREYDVK 120  
QY 121 VSAFENPYVDAIKSLMNDPGIOECYDRREYQSDSTKYLLNDLDRADPAVLPYDQDVL 180  
DB 121 VSAFENPYVDAIKSLMNDPGIOECYDRREYQSDSTKYLLNDLDRADPAVLPYDQDVL 180  
QY 181 RVAVPTGIIIEYFPDLOSIFRWDVGGQSRERKWHCFENVTSMFLVALSEYDQVYL 240  
DB 181 RVAVPTGIIIEYFPDLOSIFRWDVGGQSRERKWHCFENVTSMFLVALSEYDQVYL 240  
QY 241 ESDNENRMEESKALFRTIITYPWFONSSVILFLNKKDLLEKIMYSHLVDFPEYDGPQR 300  
DB 241 ESDNENRMEESKALFRTIITYPWFONSSVILFLNKKDLLEKIMYSHLVDFPEYDGPQR 300  
QY 301 DAQAAREFLIKMFVDLNPDSKIIYSHFTCATDTENIRFVAAYKDTIILQNLKEYNLY 359  
DB 301 DAQAAREFLIKMFVDLNPDSKIIYSHFTCATDTENIRFVAAYKDTIILQNLKEYNLY 359

RESULT 5



```
US-60-465-241-5597
; Sequence 5597, Application US/60465241
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: BEGOVICH, Ann
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CLO01468
; CURRENT APPLICATION NUMBER: US/60/465,241
; CURRENT FILING DATE: 2003-04-23
; NUMBER OF SEQ ID NOS: 258418
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 5597
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-465-241-5597

Query Match          99.4%; Score 1849; DB 7; Length 359;
Best Local Similarity 99.4%; Pred. No. 4.1e-183;
Matches 357; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLESIMACCLSEEAKEARRINDETERHVRDCKRDARRRLKLLLGTSSEKSTPIKQMR 60
DB 1 MLESIMACCLSEEAKEARRINDETERHVRDCKRDARRRLKLLLGTSSEKSTPIKQMR 60
QY 61 IIHGSGYDDEKRGKTKLVYQNIPTAMQAMIRAMDTLKIPYEEHNKAHAQLVREVDYER 120
DB 61 IIHGSGYDDEKRGKTKLVYQNIPTAMQAMIRAMDTLKIPYEEHNKAHAQLVREVDYER 120
QY 121 VSAFENPYVDAIKSLMNDPGIOECYDRRREYQSDSTKYYLNDLDRVADPAVLPQOQDVL 180
DB 121 VSAFENPYVDAIKSLMNDPGIOECYDRRREYQSDSTKYYLNDLDRVADPAVLPQOQDVL 180
QY 121 VSAFENPYVDAIKSLMNDPGIOECYDRRREYQSDSTKYYLNDLDRVADPAVLPQOQDVL 180
DB 121 VSAFENPYVDAIKSLMNDPGIOECYDRRREYQSDSTKYYLNDLDRVADPAVLPQOQDVL 180
QY 161 RVRVPTGTGIIIEYFPDLOSIVIFRMVDVGGORSERRKWIHCFENVTSIMPLVALSEYDOYLV 240
DB 161 RVRVPTGTGIIIEYFPDLOSIVIFRMVDVGGORSERRKWIHCFENVTSIMPLVALSEYDOYLV 240
QY 241 ESDNENRMEESKALFRTIITYPFONSSVILFLNKKDLLEEKIMYSHLVADYPEYDGPQR 300
DB 241 ESDNENRMEESKALFRTIITYPFONSSVILFLNKKDLLEEKIMYSHLVADYPEYDGPQR 300
QY 301 DAQAAREFLIKMFVDLNPDSKIIYSHFTCATDTENIRFVAVKDTIILQLNLKEYNLV 359
DB 301 DAQAAREFLIKMFVDLNPDSKIIYSHFTCATDTENIRFVAVKDTIILQLNLKEYNLV 359

RESULT 6
US-60-466-412-10419
; Sequence 10419, Application US/60466412
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CLO01466
; CURRENT APPLICATION NUMBER: US/60/466,412
; CURRENT FILING DATE: 2003-04-30
; NUMBER OF SEQ ID NOS: 429241
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 10419
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-466-412-10419

Query Match          99.4%; Score 1849; DB 7; Length 359;
Best Local Similarity 99.4%; Pred. No. 4.1e-183;
Matches 357; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLESIMACCLSEEAKEARRINDETERHVRDCKRDARRRLKLLLGTSSEKSTPIKQMR 60
DB 1 MLESIMACCLSEEAKEARRINDETERHVRDCKRDARRRLKLLLGTSSEKSTPIKQMR 60
```

```
QY 61 IIHGSGYDDEKRGKTKLVYQNIPTAMQAMIRAMDTLKIPYEEHNKAHAQLVREVDYER 120
DB 61 IIHGSGYDDEKRGKTKLVYQNIPTAMQAMIRAMDTLKIPYEEHNKAHAQLVREVDYER 120
QY 121 VSAFENPYVDAIKSLMNDPGIOECYDRRREYQSDSTKYYLNDLDRVADPAVLPQOQDVL 180
DB 121 VSAFENPYVDAIKSLMNDPGIOECYDRRREYQSDSTKYYLNDLDRVADPAVLPQOQDVL 180
QY 161 RVRVPTGTGIIIEYFPDLOSIVIFRMVDVGGORSERRKWIHCFENVTSIMPLVALSEYDOYLV 240
DB 161 RVRVPTGTGIIIEYFPDLOSIVIFRMVDVGGORSERRKWIHCFENVTSIMPLVALSEYDOYLV 240
QY 241 ESDNENRMEESKALFRTIITYPFONSSVILFLNKKDLLEEKIMYSHLVADYPEYDGPQR 300
DB 241 ESDNENRMEESKALFRTIITYPFONSSVILFLNKKDLLEEKIMYSHLVADYPEYDGPQR 300
QY 301 DAQAAREFLIKMFVDLNPDSKIIYSHFTCATDTENIRFVAVKDTIILQLNLKEYNLV 359
DB 301 DAQAAREFLIKMFVDLNPDSKIIYSHFTCATDTENIRFVAVKDTIILQLNLKEYNLV 359

RESULT 7
US-10-258-561-2
; Sequence 2, Application US/10258561
; GENERAL INFORMATION:
; APPLICANT: Young, Kathleen
; APPLICANT: Howland, David S.
; APPLICANT: Margulis, Karen L.
; APPLICANT: Rosenzweig-Lipson, Sharon
; APPLICANT: Cockrell, Mark Ian
; TITLE OF INVENTION: Transgenic Rat
; FILE REFERENCE: 3655/1J046US2
; CURRENT APPLICATION NUMBER: US/10/258,561
; CURRENT FILING DATE: 2002-10-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: as well as several other changes to include an EE epitope tag
US-10-258-561-2

Query Match          99.1%; Score 1844; DB 6; Length 359;
Best Local Similarity 98.9%; Pred. No. 1.4e-182;
Matches 355; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MLESIMACCLSEEAKEARRINDETERHVRDCKRDARRRLKLLLGTSSEKSTPIKQMR 60
DB 1 MLESIMACCLSEEAKEARRINDETERHVRDCKRDARRRLKLLLGTSSEKSTPIKQMR 60
QY 61 IIHGSGYDDEKRGKTKLVYQNIPTAMQAMIRAMDTLKIPYEEHNKAHAQLVREVDYER 120
DB 61 IIHGSGYDDEKRGKTKLVYQNIPTAMQAMIRAMDTLKIPYEEHNKAHAQLVREVDYER 120
QY 121 VSAFENPYVDAIKSLMNDPGIOECYDRRREYQSDSTKYYLNDLDRVADPAVLPQOQDVL 180
DB 121 VSAFENPYVDAIKSLMNDPGIOECYDRRREYQSDSTKYYLNDLDRVADPAVLPQOQDVL 180
QY 161 RVRVPTGTGIIIEYFPDLOSIVIFRMVDVGGORSERRKWIHCFENVTSIMPLVALSEYDOYLV 240
DB 161 RVRVPTGTGIIIEYFPDLOSIVIFRMVDVGGORSERRKWIHCFENVTSIMPLVALSEYDOYLV 240
QY 241 ESDNENRMEESKALFRTIITYPFONSSVILFLNKKDLLEEKIMYSHLVADYPEYDGPQR 300
DB 241 ESDNENRMEESKALFRTIITYPFONSSVILFLNKKDLLEEKIMYSHLVADYPEYDGPQR 300
QY 301 DAQAAREFLIKMFVDLNPDSKIIYSHFTCATDTENIRFVAVKDTIILQLNLKEYNLV 359
DB 301 DAQAAREFLIKMFVDLNPDSKIIYSHFTCATDTENIRFVAVKDTIILQLNLKEYNLV 359
```

Db 301 DAQAREFILKMFVDLNDPSDKIITSHTCATDTENTIRFVAANKDTIQLNLKEYNLV 359

RESULT 8  
PCT-US03-02452-13  
Sequence 13, Application PC/US0302452  
GENERAL INFORMATION:  
APPLICANT: Moore, Lisa  
APPLICANT: Kindt, Rachel  
APPLICANT: Kopczyński, Jenny  
APPLICANT: Doderstein, Stephen  
APPLICANT: Cockett, Mark  
APPLICANT: Ramanathan, Chandra  
APPLICANT: Lodge, Nicholas  
APPLICANT: Fitzgerald, Kevin  
TITLE OF INVENTION: MOLECULES THAT MODULATE G(α) ACTIVITY AND METHODS OF  
FILE REFERENCE: 5624-277-228  
CURRENT APPLICATION NUMBER: PCT/US03/02452  
CURRENT FILING DATE: 2003-01-28  
PRIOR APPLICATION NUMBER: US 60/352720  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 13  
LENGTH: 353  
TYPE: PRT  
ORGANISM: Artificial  
FEATURE:  
OTHER INFORMATION: G-protein of the invention  
PCT-US03-02452-13

Query Match 97.7%; Score 1817; DB 1; Length 353;  
Best Local Similarity 99.2%; Pred. No. 8.4e-180;  
Matches 350; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 7 MACCLSEAKRARRINDEIERHVRDRKRDARELKLILGSGSGKSTFIKQMTIHGSG 66  
Db 1 MACCLSEAKRARRINDEIERHVRDRKRDARELKLILGSGSGKSTFIKQMTIHGSG 60  
Qy 67 YSDEKRGFTKLVYQNIETAMQAMIRAMDTLKIPYKYEHNKAHQAOLREVDEKVSFEN 126  
Db 61 YSDEKRGFTKLVYQNIETAMQAMIRAMDTLKIPYKYEHNKAHQAOLREVDEKVSFEN 120  
Qy 127 PYVDAIKSLWNPDIQECYDRRREYQLSDSTKYIYLNLDRAVDAPAYLPTQODVLRVRYPT 186  
Db 121 PYVDAIKSLWNPDIQECYDRRREYQLSDSTKYIYLNLDRAVDAPAYLPTQODVLRVRYPT 180  
Qy 187 TGIIEYFPDLOSIFRNVADVGQSRERKMIHCFENYTSIMFVALSEYDQVLVESDNN 246  
Db 181 TGIIEYFPDLOSIFRNVADVGQSRERKMIHCFENYTSIMFVALSEYDQVLVESDNN 240  
Qy 247 RMESKALFRTIITYPFQNSVLLFLNKKDLLEKIMYSHLVDFPEYDGPQDAQAAR 306  
Db 241 RMESKALFRTIITYPFQNSVLLFLNKKDLLEKIMYSHLVDFPEYDGPQDAQAAR 300  
Qy 307 EFILKMFVDLNDPSDKIITSHTCATDTENTIRFVAANKDTIQLNLKEYNLV 359  
Db 301 EFILKMFVDLNDPSDKIITSHTCATDTENTIRFVAANKDTIQLNLKEYNLV 353

RESULT 9  
US-10-352-843-13  
Sequence 13, Application US/10352843  
GENERAL INFORMATION:  
APPLICANT: Moore, Lisa  
APPLICANT: Kindt, Rachel  
APPLICANT: Kopczyński, Jenny  
APPLICANT: Doderstein, Stephen  
APPLICANT: Cockett, Mark  
APPLICANT: Ramanathan, Chandra  
APPLICANT: Lodge, Nicholas

APPLICANT: Fitzgerald, Kevin  
APPLICANT: Stouch, Terry  
TITLE OF INVENTION: MOLECULES THAT MODULATE G(α) ACTIVITY AND METHODS OF  
FILE REFERENCE: 5624-277-999  
CURRENT APPLICATION NUMBER: US/10/352,843  
CURRENT FILING DATE: 2003-01-27  
PRIOR APPLICATION NUMBER: US 60/352720  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 13  
LENGTH: 353  
TYPE: PRT  
ORGANISM: Artificial  
FEATURE:  
OTHER INFORMATION: G-protein of the invention  
US-10-352-843-13

Query Match 97.7%; Score 1817; DB 6; Length 353;  
Best Local Similarity 99.2%; Pred. No. 8.4e-180;  
Matches 350; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy\* 7 MACCLSEAKRARRINDEIERHVRDRKRDARELKLILGSGSGKSTFIKQMTIHGSG 66  
Db 1 MACCLSEAKRARRINDEIERHVRDRKRDARELKLILGSGSGKSTFIKQMTIHGSG 60  
Qy 67 YSDEKRGFTKLVYQNIETAMQAMIRAMDTLKIPYKYEHNKAHQAOLREVDEKVSFEN 126  
Db 61 YSDEKRGFTKLVYQNIETAMQAMIRAMDTLKIPYKYEHNKAHQAOLREVDEKVSFEN 120  
Qy 127 PYVDAIKSLWNPDIQECYDRRREYQLSDSTKYIYLNLDRAVDAPAYLPTQODVLRVRYPT 186  
Db 121 PYVDAIKSLWNPDIQECYDRRREYQLSDSTKYIYLNLDRAVDAPAYLPTQODVLRVRYPT 180  
Qy 187 TGIIEYFPDLOSIFRNVADVGQSRERKMIHCFENYTSIMFVALSEYDQVLVESDNN 246  
Db 181 TGIIEYFPDLOSIFRNVADVGQSRERKMIHCFENYTSIMFVALSEYDQVLVESDNN 240  
Qy 247 RMESKALFRTIITYPFQNSVLLFLNKKDLLEKIMYSHLVDFPEYDGPQDAQAAR 306  
Db 241 RMESKALFRTIITYPFQNSVLLFLNKKDLLEKIMYSHLVDFPEYDGPQDAQAAR 300  
Qy 307 EFILKMFVDLNDPSDKIITSHTCATDTENTIRFVAANKDTIQLNLKEYNLV 359  
Db 301 EFILKMFVDLNDPSDKIITSHTCATDTENTIRFVAANKDTIQLNLKEYNLV 353

RESULT 10  
US-09-724-676-49632  
Sequence 49632, Application US/09724676  
GENERAL INFORMATION:  
APPLICANT: Compugen LTD  
TITLE OF INVENTION: Variants of alternative splicing  
FILE REFERENCE: 129181.4 Compugen  
CURRENT APPLICATION NUMBER: US/09/724,676  
CURRENT FILING DATE: 2000-11-28  
NUMBER OF SEQ ID NOS: 97222  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 49632  
LENGTH: 387  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-724-676-49632

Query Match 90.7%; Score 1687; DB 5; Length 387;  
Best Local Similarity 89.4%; Pred. No. 3e-166;  
Matches 321; Conservative 23; Mismatches 15; Indels 0; Gaps 0;

Qy 1 MTLESIMACCLSEAKRARRINDEIERHVRDRKRDARELKLILGSGSGKSTFIKQMR 60  
Db 1 MTLESIMACCLSEAKRARRINDEIERHVRDRKRDARELKLILGSGSGKSTFIKQMR 60

QY 61 IIHSGYSDKRGFTKLVYONIFETAMQAMIRAMDTLKIPYKEHNKAHQAQVREVDYER 120  
; FILE REFERENCE: C000307  
; CURRENT APPLICATION NUMBER: US/09/949, 016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 7853  
; LENGTH: 351  
; TYPE: PRF  
; ORGANISM: Human  
US-09-949-016-7853

QY 241 ESDNENRMEESKALFRTIITYPWFONSSVILFLNKKDLLEKIMYSHLVDYFPEYDQPOR 300  
; FILE REFERENCE: C000307  
; CURRENT APPLICATION NUMBER: US/09/949, 016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 7853  
; LENGTH: 351  
; TYPE: PRF  
; ORGANISM: Human  
US-09-949-016-7853

QY 241 ESDNENRMEESKALFRTIITYPWFONSSVILFLNKKDLLEKIMYSHLVDYFPEYDQPOR 300  
; FILE REFERENCE: C000307  
; CURRENT APPLICATION NUMBER: US/09/949, 016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 7853  
; LENGTH: 351  
; TYPE: PRF  
; ORGANISM: Human  
US-09-949-016-7853

QY 301 DAQAAREFLIKMFVDLNDPDSKIIYSHFTCATDTENIRFVAAYKDTIIOQLNKEYNLY 359  
; FILE REFERENCE: C000307  
; CURRENT APPLICATION NUMBER: US/09/949, 016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 7853  
; LENGTH: 351  
; TYPE: PRF  
; ORGANISM: Human  
US-09-949-016-7853

RESULT 11  
US-09-724-676A-49632  
; Sequence 49632, Application US/09724676A  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724, 676A  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 49632  
; LENGTH: 387  
; TYPE: PRF  
; ORGANISM: Homo sapiens  
US-09-724-676A-49632

Query Match 90.7%; Score 1687; DB 5; Length 387;  
Best Local Similarity 89.4%; Pred. No. 3e-166;  
Matches 321; Conservative 23; Mismatches 15; Indels 0; Gaps 0;

QY 1 MLESIYACCLSEBAKRIINDEIERHVRDRKRDARRELKLLIGTGESGKSTFIKQMR 60  
; FILE REFERENCE: C000307  
; CURRENT APPLICATION NUMBER: US/09/724, 676A  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 49632  
; LENGTH: 387  
; TYPE: PRF  
; ORGANISM: Homo sapiens  
US-09-724-676A-49632

QY 61 IIHSGYSDKRGFTKLVYONIFETAMQAMIRAMDTLKIPYKEHNKAHQAQVREVDYER 120  
; FILE REFERENCE: C000307  
; CURRENT APPLICATION NUMBER: US/09/949, 016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 7853  
; LENGTH: 351  
; TYPE: PRF  
; ORGANISM: Human  
US-09-949-016-7853

QY 121 VSAFENPYDAIKSLMNDPGIOECYDRREYQSLDSTKYLLNDLDRVADPAVYLPQODVL 180  
; FILE REFERENCE: C000307  
; CURRENT APPLICATION NUMBER: US/09/949, 016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 7853  
; LENGTH: 351  
; TYPE: PRF  
; ORGANISM: Human  
US-09-949-016-7853

QY 181 RVRVPTGTIIEYFEDLNIIFRWVDVGQSRERKWHICEFNVSIMFLVASEYDQVLY 240  
; FILE REFERENCE: C000307  
; CURRENT APPLICATION NUMBER: US/09/949, 016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 7853  
; LENGTH: 351  
; TYPE: PRF  
; ORGANISM: Human  
US-09-949-016-7853

QY 241 ESDNENRMEESKALFRTIITYPWFONSSVILFLNKKDLLEKIMYSHLVDYFPEYDQPOR 300  
; FILE REFERENCE: C000307  
; CURRENT APPLICATION NUMBER: US/09/949, 016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 7853  
; LENGTH: 351  
; TYPE: PRF  
; ORGANISM: Human  
US-09-949-016-7853

QY 301 DAQAAREFLIKMFVDLNDPDSKIIYSHFTCATDTENIRFVAAYKDTIIOQLNKEYNLY 359  
; FILE REFERENCE: C000307  
; CURRENT APPLICATION NUMBER: US/09/949, 016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 7853  
; LENGTH: 351  
; TYPE: PRF  
; ORGANISM: Human  
US-09-949-016-7853

RESULT 12  
US-09-949-016-7853  
; Sequence 7853, Application US/09949016  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

QY 9 CCISEAKKARINDEIERHVRDRKRDARRELKLLIGTGESGKSTFIKQMR11HSGYS 68  
; FILE REFERENCE: C000307  
; CURRENT APPLICATION NUMBER: US/09/949, 016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 7853  
; LENGTH: 351  
; TYPE: PRF  
; ORGANISM: Human  
US-09-949-016-7853

QY 1 CCISDEYKESKRINAEIEKQLRDRKRDARRELKLLIGTGESGKSTFIKQMR11HSGYS 60  
; FILE REFERENCE: C000307  
; CURRENT APPLICATION NUMBER: US/09/949, 016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 7853  
; LENGTH: 351  
; TYPE: PRF  
; ORGANISM: Human  
US-09-949-016-7853

QY 69 DEKRGFTKLVYONIFETAMQAMIRAMDTLKIPYKEHNKAHQAQVREVDYER 128  
; FILE REFERENCE: C000307  
; CURRENT APPLICATION NUMBER: US/09/949, 016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 7853  
; LENGTH: 351  
; TYPE: PRF  
; ORGANISM: Human  
US-09-949-016-7853

QY 129 VDAIKSLMNDPGIOECYDRREYQSLDSTKYLLNDLDRVADPAVYLPQODVL 188  
; FILE REFERENCE: C000307  
; CURRENT APPLICATION NUMBER: US/09/949, 016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 7853  
; LENGTH: 351  
; TYPE: PRF  
; ORGANISM: Human  
US-09-949-016-7853

QY 189 IIEYFEDLNIIFRWVDVGQSRERKWHICEFNVSIMFLVASEYDQVLY 248  
; FILE REFERENCE: C000307  
; CURRENT APPLICATION NUMBER: US/09/949, 016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 7853  
; LENGTH: 351  
; TYPE: PRF  
; ORGANISM: Human  
US-09-949-016-7853

QY 249 EESKALFRTIITYPWFONSSVILFLNKKDLLEKIMYSHLVDYFPEYDQPOR 308  
; FILE REFERENCE: C000307  
; CURRENT APPLICATION NUMBER: US/09/949, 016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 7853  
; LENGTH: 351  
; TYPE: PRF  
; ORGANISM: Human  
US-09-949-016-7853

QY 309 ILKMFVDLNDPDSKIIYSHFTCATDTENIRFVAAYKDTIIOQLNKEYNLY 359  
; FILE REFERENCE: C000307  
; CURRENT APPLICATION NUMBER: US/09/949, 016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 7853  
; LENGTH: 351  
; TYPE: PRF  
; ORGANISM: Human  
US-09-949-016-7853

Query Match 82.5%; Score 1535; DB 6; Length 351;  
Best Local Similarity 82.1%; Pred. No. 1.5e-150;  
Matches 288; Conservative 35; Mismatches 28; Indels 0; Gaps 0;

RESULT 13  
US-10-326-896-2  
; Sequence 2, Application US/10326896  
; GENERAL INFORMATION:  
; APPLICANT: Keith D. Allen  
; TITLE OF INVENTION: GNA14 G-Protein Signaling Protein  
; TITLE OF INVENTION: Subunit Gene Disruptions, Compositions and Methods Related  
; FILE REFERENCE: R-696  
; CURRENT APPLICATION NUMBER: US/10/326, 896  
; CURRENT FILING DATE: 2002-12-20  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 355  
; TYPE: PRF  
; ORGANISM: Mus musculus  
US-10-326-896-2

QY 9 CCISEAKKARINDEIERHVRDRKRDARRELKLLIGTGESGKSTFIKQMR11HSGYS 68  
; FILE REFERENCE: C000307  
; CURRENT APPLICATION NUMBER: US/09/949, 016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 7853  
; LENGTH: 351  
; TYPE: PRF  
; ORGANISM: Human  
US-09-949-016-7853

Db 5 CCLSAEESQISAEIERHVRDKDARRELKLLLGESGSKSTFIKOMRIIHSGS 64  
Qy 69 DEDKRGFTKLVYONIFITAMQAMIRAMDTLKIPYKYEHNKAHQAOLVREVDVEKVSAPENPY 128  
Db 65 DEBRKGTFLVYONIFITAMQAMIRAMDTLKIPYKYEHNKAHQAOLVREVDVEKVSAPENPY 124  
Qy 129 VDAIKSLMNDPGIOECYDRRREYQOLSDSTKYVNLNDLRVADPAYLPTQODVLRVPTTG 188  
Db 125 VAAIKLMDLPGIOECYDRRREYQOLSDSTKYVNLNDLRVADPAYLPTQODVLRVPTTG 184  
Qy 189 IIEYPPDLQSVIRAWDVGGORSEBRKWHCFENVTSMIFVALSEYDOVLVSDNENRM 248  
Db 185 IIEYPPDLQSVIRAWDVGGORSEBRKWHCFENVTSMIFVALSEYDOVLVSDNENRM 244  
Qy 249 EESKALFRTIITYPWFONSSVILFLNKKDLLEEKIMYSHLVDFPEYDGPORDAQAARF 308  
Db 245 EESKALFRTIITYPWFONSSVILFLNKKDLLEEKIMYSHLVDFPEYDGPORDAQAARF 304  
Qy 309 ILKMFVDLNPDSKIIYSHFTCATDTENIRFVFAAVKDTILOHNLKEYNLV 359  
Db 305 ILKMFVDLNPDSKIIYSHFTCATDTENIRFVFAAVKDTILOHNLKEYNLV 355

## RESULT 14

PCT-US03-02452-12  
Sequence 12, Application PC/TUS0302452  
GENERAL INFORMATION:  
APPLICANT: Moore, Lisa  
APPLICANT: Kindt, Rachel  
APPLICANT: Koczynski, Jenny  
APPLICANT: Dobersztajn, Stephen  
APPLICANT: Cocke, Mark  
APPLICANT: Ramanathan, Chandra  
APPLICANT: Fitzgerald, Kevin  
APPLICANT: Stouch, Terry  
TITLE OF INVENTION: MOLECULES THAT MODULATE G(1/2) ACTIVITY AND METHODS OF  
FILE REFERENCE: 5624-277-228  
CURRENT FILING DATE: 2003-01-28  
PRIOR APPLICATION NUMBER: US 60/352720  
PRIOR FILING DATE: 2002-01-28  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 12  
LENGTH: 355  
TYPE: PRT  
ORGANISM: Caenorhabditis elegans  
PCT-US03-02452-12

Query Match 81.3%; Score 1513; DB 1; Length 355;  
Best Local Similarity 82.0%; Pred. No. 3e-148;  
Matches 291; Conservative 26; Mismatches 36; Indels 2; Gaps 1;

Qy 7 MACCSEEAKEARINDEIERHVRDKDARRELKLLLGESGSKSTFIKOMRIIHSGS 66  
Db 1 MACCSEEAKEARINDEIERHVRDKDARRELKLLLGESGSKSTFIKOMRIIHSGS 60  
Qy 67 YSDEKRGFTKLVYONIFITAMQAMIRAMDTLKIPYKYEHNKAHQAOLVREVDVEKVSAP 124  
Db 61 YSDEKRGFTKLVYONIFITAMQAMIRAMDTLKIPYKYEHNKAHQAOLVREVDVEKVSAP 120  
Qy 125 ENPYDAIKSLMNDPGIOECYDRRREYQOLSDSTKYVNLNDLRVADPAYLPTQODVLRV 184  
Db 121 EEPYDAIKSLMNDPGIOECYDRRREYQOLSDSTKYVNLNDLRVADPAYLPTQODVLRV 180  
Qy 185 PTTGIEYPPDLQSVIRAWDVGGORSEBRKWHCFENVTSMIFVALSEYDOVLVSDNENRM 244  
Db 181 PTTGIEYPPDLQSVIRAWDVGGORSEBRKWHCFENVTSMIFVALSEYDOVLVSDNENRM 240  
Qy 245 ENRMESKALFRTIITYPWFONSSVILFLNKKDLLEEKIMYSHLVDFPEYDGPORDAQA 304  
Db 245 ENRMESKALFRTIITYPWFONSSVILFLNKKDLLEEKIMYSHLVDFPEYDGPORDAQA 304

Db 241 ENRMESKALFRTIITYPWFONSSVILFLNKKDLLEEKIMYSHLVDFPEYDGPORDAQA 300  
Qy 305 AREFIIKMFVDLNPDSKIIYSHFTCATDTENIRFVFAAVKDTILOHNLKEYNLV 359  
Db 301 AREFIIKMFVDLNPDSKIIYSHFTCATDTENIRFVFAAVKDTILOHNLKEYNLV 355

## RESULT 15

US-10-369-493-4982  
Sequence 4982, Application US/10369493  
GENERAL INFORMATION:  
APPLICANT: Cao, Yongwei  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Goldman, Barry S.  
APPLICANT: Chen, Xianfeng  
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
FILE REFERENCE: 38-10(52052)B  
CURRENT FILING DATE: US/10/369,493  
CURRENT FILING DATE: 2003-02-28  
PRIOR APPLICATION NUMBER: US 60/360,039  
PRIOR FILING DATE: 2002-02-21  
NUMBER OF SEQ ID NOS: 47374  
SEQ ID NO 4982  
LENGTH: 355  
TYPE: PRT  
ORGANISM: Caenorhabditis elegans  
US-10-369-493-4982

Query Match 81.3%; Score 1513; DB 6; Length 355;  
Best Local Similarity 82.0%; Pred. No. 3e-148;  
Matches 291; Conservative 26; Mismatches 36; Indels 2; Gaps 1;

Qy 7 MACCSEEAKEARINDEIERHVRDKDARRELKLLLGESGSKSTFIKOMRIIHSGS 66  
Db 1 MACCSEEAKEARINDEIERHVRDKDARRELKLLLGESGSKSTFIKOMRIIHSGS 60  
Qy 67 YSDEKRGFTKLVYONIFITAMQAMIRAMDTLKIPYKYEHNKAHQAOLVREVDVEKVSAP 124  
Db 61 YSDEKRGFTKLVYONIFITAMQAMIRAMDTLKIPYKYEHNKAHQAOLVREVDVEKVSAP 120  
Qy 125 ENPYDAIKSLMNDPGIOECYDRRREYQOLSDSTKYVNLNDLRVADPAYLPTQODVLRV 184  
Db 121 EEPYDAIKSLMNDPGIOECYDRRREYQOLSDSTKYVNLNDLRVADPAYLPTQODVLRV 180  
Qy 185 PTTGIEYPPDLQSVIRAWDVGGORSEBRKWHCFENVTSMIFVALSEYDOVLVSDNENRM 244  
Db 181 PTTGIEYPPDLQSVIRAWDVGGORSEBRKWHCFENVTSMIFVALSEYDOVLVSDNENRM 240  
Qy 245 ENRMESKALFRTIITYPWFONSSVILFLNKKDLLEEKIMYSHLVDFPEYDGPORDAQA 304  
Db 241 ENRMESKALFRTIITYPWFONSSVILFLNKKDLLEEKIMYSHLVDFPEYDGPORDAQA 300  
Qy 305 AREFIIKMFVDLNPDSKIIYSHFTCATDTENIRFVFAAVKDTILOHNLKEYNLV 359  
Db 301 AREFIIKMFVDLNPDSKIIYSHFTCATDTENIRFVFAAVKDTILOHNLKEYNLV 355

Search completed: June 30, 2003, 16:25:31  
Job time : 52 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2003, 16:09:36 ; Search time 39 Seconds

(without alignments)  
1226.589 Million cell updates/sec

Title: US-09-899-295-2

Perfect score: 1860

Sequence: 1 MTELSIMACLSSEAKEARR.....VFAAVKPTIQLNKEYNLV 359

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA2001.DAT.\*  
23: /SIDS2/gcgdata/geneseq/genesqp-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1860	100.0	359	23	AAM48816 Murine G-protein 6
2	1849	99.4	359	23	ABG68610 Human G-protein al
3	1846	99.2	359	23	ABG68584 Mouse G-protein al
4	1844	99.1	359	23	ABG68598 Human G-protein al
5	1827	98.2	1276	20	AAV49127 Human p115R2*Gq15
6	1827	98.2	1276	23	AAO15093 Human p115R2*Gq15
7	1827	98.2	1394	20	AAV49129 Human p115R2*Gq15
8	1827	98.2	1394	23	AAO15095 Human p115R2*Gq15
9	1827	98.2	1397	20	AAV49134 Human p115R2*Gq15
10	1827	98.2	1397	23	AAO15100 Human p115R2*Gq15

11	1827	98.2	1402	23	AAO15105 Human p115R2*Gq15
12	1827	98.2	1418	20	AAV49131 Human p115R2*Gq15
13	1827	98.2	1418	23	AAO15097 Human p115R2*Gq15
14	1827	98.2	1421	23	AAO15103 Human p115R2*Gq15
15	1827	98.2	1422	23	AAO15102 Human p115R2*Gq15
16	1823	98.0	1323	20	AAV49133 Human p115R2*Gq15
17	1823	98.0	1323	23	AAO15099 Human p115R2*Gq15
18	1822	98.0	359	23	ABG68592 Human p115R2*Gq15
19	1819	97.8	359	23	ABG68585 Human p115R2*Gq15
20	1819	97.8	359	23	AAV49125 Human p115R2*Gq15
21	1818	97.7	359	23	ABG68604 Human p115R2*Gq15
22	1817	97.7	353	23	ABG68597 Human p115R2*Gq15
23	1817	97.7	353	23	ABG68599 Human p115R2*Gq15
24	1817	97.7	353	23	ABG68591 Human p115R2*Gq15
25	1816	97.6	1303	20	AAV49132 Human p115R2*Gq15
26	1815	97.6	1303	23	AAO15098 Human p115R2*Gq15
27	1815	97.6	359	23	ABG68594 Human p115R2*Gq15
28	1814	97.5	353	23	AAV49131 Human p115R2*Gq15
29	1814	97.5	353	23	AAV49131 Human p115R2*Gq15
30	1813	97.5	359	23	ABG68608 Human p115R2*Gq15
31	1811	97.4	359	23	ABG68586 Human p115R2*Gq15
32	1810	97.3	353	23	AAM48817 Human p115R2*Gq15
33	1807	97.2	359	23	ABG68603 Human p115R2*Gq15
34	1806	97.1	359	23	ABG68609 Human p115R2*Gq15
35	1804	97.0	359	23	ABG68593 Human p115R2*Gq15
36	1803	96.9	359	23	ABG68605 Human p115R2*Gq15
37	1800	96.8	359	23	ABG68595 Human p115R2*Gq15
38	1796	96.6	359	23	ABG68606 Human p115R2*Gq15
39	1795	96.5	353	23	ABG68591 Human p115R2*Gq15
40	1793	96.4	359	23	ABG68596 Human p115R2*Gq15
41	1791	96.3	353	23	ABG68600 Human p115R2*Gq15
42	1788	96.1	353	23	ABG68588 Human p115R2*Gq15
43	1784	95.9	353	23	ABG68587 Human p115R2*Gq15
44	1779	95.6	353	23	ABG68607 Human p115R2*Gq15
45	1776	95.5	353	23	ABG68597 Human p115R2*Gq15

# ALIGNMENTS

RESULT 1	
AAM48816	
ID	AAM48816 standard; Protein; 359 AA.
XX	
AC	AAM48816;
XX	
DT	24-APR-2002 (first entry)
XX	
DE	Murine G-protein Gq14myr.
XX	
KW	Mouse; G-protein coupled receptor modulator; G-protein Gq14myr;
KW	signal transduction.
XX	
OS	Mus musculus.
XX	
PN	WO200204665-A2.
XX	
PD	17-JAN-2002.
XX	
PF	05-JUL-2001; 2001WO-EP07667.
XX	
PR	08-JUL-2000; 2000DE-1033353.
XX	
PA	(AVET) AVANTIS PHARMA DEUT GMBH.
XX	
PI	Kostenis E;
XX	
DR	WPI; 2002-148182/19.
XX	
DR	N-PSDB; ABA97518.
XX	
PT	Identifying compounds that modify activity of signal transduction
PT	pathways, useful potentially as therapeutic agents, by screening with
PT	cells that contain hybrid G proteins

XX Claim 22; Page 24-25; 34pp; German.  
PS  
XX  
CC The present invention relates to a method of identifying compounds which  
CC are capable of modifying the activity of a signal transduction pathway  
CC which is dependent upon a G-protein coupled receptor. These compounds may  
CC include the novel G-proteins Gq14myr (shown here), Gq15myr, Gq14, Gq55  
CC and Galphal6.  
XX  
SQ Sequence 359 AA;  
Query Match 100.0%; Score 1860; DB 23; Length 359;  
Best Local Similarity 100.0%; Pred. No. 3.6e-176; Indels 0; Gaps 0;  
Matches 359; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MTEESIMACCLSEBAKEARRINDEIERHVRDRKRDARRELKLLLTGSGSKSTFIKMR 60  
DB 1 MTEESIMACCLSEBAKEARRINDEIERHVRDRKRDARRELKLLLTGSGSKSTFIKMR 60  
QY 61 IIHSGSYSDCKRGFTKLVYONITFTAMQAMIRAMDTLKIPYKYEHNKAHQVREVDYER 120  
DB 61 IIHSGSYSDCKRGFTKLVYONITFTAMQAMIRAMDTLKIPYKYEHNKAHQVREVDYER 120  
QY 121 VSAFENFYVDAIKSLMNDPGIOECYDRRREYQLSDSKYKYLNDLRVADPAVYLPDQDVL 180  
DB 121 VSAFENFYVDAIKSLMNDPGIOECYDRRREYQLSDSKYKYLNDLRVADPAVYLPDQDVL 180  
QY 181 RVRVPTGTIIEYPPDLOSIFRMVDVGQSRERKRWIHCENWTSIMFLVALSEYDQVL 240  
DB 181 RVRVPTGTIIEYPPDLOSIFRMVDVGQSRERKRWIHCENWTSIMFLVALSEYDQVL 240  
QY 241 ESDNENMESKALFRITITYPWFONSSVILFLKKDLLEKIMYSHLVDFPEYDGPOR 300  
DB 241 ESDNENMESKALFRITITYPWFONSSVILFLKKDLLEKIMYSHLVDFPEYDGPOR 300  
QY 301 DAQAAREFIKMFVDLPDSDKIITSHFTCATDTENTIRFVFAAVKDTIILQNLKEYNLV 359  
DB 301 DAQAAREFIKMFVDLPDSDKIITSHFTCATDTENTIRFVFAAVKDTIILQNLKEYNLV 359  
RESULT 2  
ABG68610  
ID ABG68610 standard; Protein: 359 AA.  
XX  
AC ABG68610;  
XX  
DT 07-OCT-2002 (first entry)  
XX  
DE Human G protein alpha sub-unit q family #2.  
XX  
KW G-g protein; sensory signaling; chemoreceptor; tastant; olfactant;  
KW pheromone; G protein alpha sub-unit; q family; G alpha q.  
XX  
OS Homo sapiens.  
XX  
PN WO200236622-A2.  
XX  
PD 10-MAY-2002.  
XX  
PF 24-OCT-2001; 2001WO-US32619.  
XX  
PR 30-OCT-2000; 2000US-24370P.  
XX  
PA (SENO-) SENOMYX INC.  
XX  
PI Yao Y, Xu H;  
XX  
DR WPI; 2002-519234/55.  
PT New G-alpha-q protein variants, useful for analyzing and discovering  
PT agonists or antagonists of chemoreceptors, such as G protein coupled  
PT receptors involved in sensing of tastants, olfactants or pheromones  
XX

PS Disclosure; Fig 2; 32pp; English.  
XX  
CC The invention describes an isolated variant of a G-g protein, which  
CC exhibits increased promiscuity relative to the corresponding G-g protein.  
CC The variant is used to identify a compound that modulates sensory  
CC signaling in sensory cells and to identify a compound that interacts  
CC with the G-g variant protein. The G-g protein variant is useful for  
CC analyzing and discovering agonists or antagonists of chemoreceptors, such  
CC as G protein coupled receptors involved in sensing of tastants,  
CC olfactants or pheromones. This is the amino acid sequence of a G protein  
CC alpha sub-unit q family (G alpha q) protein that can functionally couple  
CC to sensory cell receptors such as taste GPCR's (G protein-coupled  
CC receptors) and olfactory GPCR's in an overly promiscuous manner.  
XX  
SQ Sequence 359 AA;  
Query Match 99.4%; Score 1849; DB 23; Length 359;  
Best Local Similarity 99.4%; Pred. No. 4.4e-175; Indels 0; Gaps 0;  
Matches 357; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MTEESIMACCLSEBAKEARRINDEIERHVRDRKRDARRELKLLLTGSGSKSTFIKMR 60  
DB 1 MTEESIMACCLSEBAKEARRINDEIERHVRDRKRDARRELKLLLTGSGSKSTFIKMR 60  
QY 61 IIHSGSYSDCKRGFTKLVYONITFTAMQAMIRAMDTLKIPYKYEHNKAHQVREVDYER 120  
DB 61 IIHSGSYSDCKRGFTKLVYONITFTAMQAMIRAMDTLKIPYKYEHNKAHQVREVDYER 120  
QY 121 VSAFENFYVDAIKSLMNDPGIOECYDRRREYQLSDSKYKYLNDLRVADPAVYLPDQDVL 180  
DB 121 VSAFENFYVDAIKSLMNDPGIOECYDRRREYQLSDSKYKYLNDLRVADPAVYLPDQDVL 180  
QY 181 RVRVPTGTIIEYPPDLOSIFRMVDVGQSRERKRWIHCENWTSIMFLVALSEYDQVL 240  
DB 181 RVRVPTGTIIEYPPDLOSIFRMVDVGQSRERKRWIHCENWTSIMFLVALSEYDQVL 240  
QY 241 ESDNENMESKALFRITITYPWFONSSVILFLKKDLLEKIMYSHLVDFPEYDGPOR 300  
DB 241 ESDNENMESKALFRITITYPWFONSSVILFLKKDLLEKIMYSHLVDFPEYDGPOR 300  
QY 301 DAQAAREFIKMFVDLPDSDKIITSHFTCATDTENTIRFVFAAVKDTIILQNLKEYNLV 359  
DB 301 DAQAAREFIKMFVDLPDSDKIITSHFTCATDTENTIRFVFAAVKDTIILQNLKEYNLV 359  
RESULT 3  
ABG68584  
ID ABG68584 standard; Protein: 359 AA.  
XX  
AC ABG68584;  
XX  
DT 07-OCT-2002 (first entry)  
XX  
DE Mouse G protein alpha sub-unit q family.  
XX  
KW G-g protein; sensory signaling; chemoreceptor; tastant; olfactant;  
KW pheromone; G protein alpha sub-unit; q family; G alpha q.  
XX  
OS Mus musculus.  
XX  
PN WO200236622-A2.  
XX  
PD 10-MAY-2002.  
XX  
PF 24-OCT-2001; 2001WO-US32619.  
XX  
PR 30-OCT-2000; 2000US-24370P.  
XX  
PA (SENO-) SENOMYX INC.  
XX  
PI Yao Y, Xu H;  
XX  
DR WPI; 2002-519234/55.

XX New G-alpha-q protein variants, useful for analyzing and discovering  
PT agonists or antagonists of chemoreceptors, such as G protein coupled  
PT receptors involved in sensing of tastants, olfactants or pheromones  
XX  
PS Claim 13; Fig 1; 32pp; English.  
XX  
CC The invention describes an isolated variant of a G-q protein, which  
CC exhibits increased promiscuity relative to the corresponding G-q protein.  
CC The variant is used to identify a compound that modulates sensory  
CC signaling in sensory cells and to identify a compound that interacts  
CC with the G-q variant protein. The G-q protein variant is useful for  
CC analysing and discovering agonists or antagonists of chemoreceptors, such  
CC as G protein coupled receptors involved in sensing of tastants,  
CC olfactants or pheromones. This is the amino acid sequence of a G protein  
CC alpha sub-unit q family (G alpha q) protein that can functionally couple  
CC to sensory cell receptors such as taste GPCR's (G protein-coupled  
CC receptors) and olfactory GPCR's in an overly promiscuous manner.  
XX  
SQ Sequence 359 AA;

Query Match 99.2%; Score 1846; DB 23; Length 359;  
Best Local Similarity 99.2%; Pred. No. 8.8e-175;  
Matches 356; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MTELSIMACCLSEBAKARRINDEIERHVRDRKRDARRELKLLLTGSGSKSTFIKQMR 60  
DB 1 MTELSIMACCLSEBAKARRINDEIERHVRDRKRDARRELKLLLTGSGSKSTFIKQMR 60  
QY 61 ITHSGYSDEDKRGFTLVYONITFAMQAMIRAMDITKIPKYEHNKAHQVREYVEK 120  
DB 61 ITHSGYSDEDKRGFTLVYONITFAMQAMIRAMDITKIPKYEHNKAHQVREYVEK 120  
QY 121 VSAFENYVDAIKSLMDPGIOECYDRRREYQLSDSKRYLLNDIRADPAVYLPYQDVL 180  
DB 121 VSAFENYVDAIKSLMDPGIOECYDRRREYQLSDSKRYLLNDIRADPAVYLPYQDVL 180  
QY 181 RVRVPTTGIIIEYPPDLSVIFRMVDVGQSRERKWIHCENVTSIMFLVALSEYDVLV 240  
DB 181 RVRVPTTGIIIEYPPDLSVIFRMVDVGQSRERKWIHCENVTSIMFLVALSEYDVLV 240  
QY 241 ESDNENMESKALFRITITYPWONSIVILFKKKDLLEKIMYSHLVDFEYDGPQR 300  
DB 241 ESDNENMESKALFRITITYPWONSIVILFKKKDLLEKIMYSHLVDFEYDGPQR 300  
QY 301 DAQAREFIILKMFVDLPDSDKIIYSHFTCATDTENTIRFVPAVKDTILLQNLKEYNLV 359  
DB 301 DAQAREFIILKMFVDLPDSDKIIYSHFTCATDTENTIRFVPAVKDTILLQNLKEYNLV 359

RESULT 4  
ABG68598  
ID ABG68598 standard; Protein: 359 AA.

XX ABG68598;

XX 07-OCT-2002 (first entry)

XX Human G protein alpha sub-unit q family #1.

XX G-q protein; sensory signaling; chemoreceptor; tastant; olfactant;  
KW pheromone; G protein alpha sub-unit; q family; G alpha q.

XX Homo sapiens.

XX MO200236622-A2.

XX 10-MAY-2002.

XX 24-OCT-2001; 2001WO-US32619.

XX 30-OCT-2000; 2000US-243770P.

XX

PA (SENO-) SENOMYX INC.  
XX  
PI Yao Y, Xu H;  
XX  
XX WPI: 2002-519234/55.

XX New G-alpha-q protein variants, useful for analyzing and discovering  
PT agonists or antagonists of chemoreceptors, such as G protein coupled  
PT receptors involved in sensing of tastants, olfactants or pheromones  
XX  
PS Claim 13; Fig 1; 32pp; English.

XX The invention describes an isolated variant of a G-q protein, which  
XX exhibits increased promiscuity relative to the corresponding G-q protein.  
XX The variant is used to identify a compound that modulates sensory  
XX signaling in sensory cells and to identify a compound that interacts  
XX with the G-q variant protein. The G-q protein variant is useful for  
XX analysing and discovering agonists or antagonists of chemoreceptors, such  
XX as G protein coupled receptors involved in sensing of tastants,  
XX olfactants or pheromones. This is the amino acid sequence of a G protein  
XX alpha sub-unit q family (G alpha q) protein that can functionally couple  
XX to sensory cell receptors such as taste GPCR's (G protein-coupled  
XX receptors) and olfactory GPCR's in an overly promiscuous manner.

Query Match 99.1%; Score 1844; DB 23; Length 359;  
Best Local Similarity 99.2%; Pred. No. 1.4e-174;  
Matches 356; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MTELSIMACCLSEBAKARRINDEIERHVRDRKRDARRELKLLLTGSGSKSTFIKQMR 60  
DB 1 MTELSIMACCLSEBAKARRINDEIERHVRDRKRDARRELKLLLTGSGSKSTFIKQMR 60  
QY 61 ITHSGYSDEDKRGFTLVYONITFAMQAMIRAMDITKIPKYEHNKAHQVREYVEK 120  
DB 61 ITHSGYSDEDKRGFTLVYONITFAMQAMIRAMDITKIPKYEHNKAHQVREYVEK 120  
QY 121 VSAFENYVDAIKSLMDPGIOECYDRRREYQLSDSKRYLLNDIRADPAVYLPYQDVL 180  
DB 121 VSAFENYVDAIKSLMDPGIOECYDRRREYQLSDSKRYLLNDIRADPAVYLPYQDVL 180  
QY 181 RVRVPTTGIIIEYPPDLSVIFRMVDVGQSRERKWIHCENVTSIMFLVALSEYDVLV 240  
DB 181 RVRVPTTGIIIEYPPDLSVIFRMVDVGQSRERKWIHCENVTSIMFLVALSEYDVLV 240  
QY 241 ESDNENMESKALFRITITYPWONSIVILFKKKDLLEKIMYSHLVDFEYDGPQR 300  
DB 241 ESDNENMESKALFRITITYPWONSIVILFKKKDLLEKIMYSHLVDFEYDGPQR 300  
QY 301 DAQAREFIILKMFVDLPDSDKIIYSHFTCATDTENTIRFVPAVKDTILLQNLKEYNLV 359  
DB 301 DAQAREFIILKMFVDLPDSDKIIYSHFTCATDTENTIRFVPAVKDTILLQNLKEYNLV 359

RESULT 5  
AAV49127  
ID AAV49127 standard; Protein: 1276 AA.

XX AAV49127;

XX 07-JAN-2000 (first entry)

XX pHcAR/hmgIur2\*Gq15 fusion construct protein sequence.

XX G-protein fusion receptor; Car; calcium receptor; Glur; head injury;  
KW metabotropic glutamate receptor; GABA(B); glutamic receptor; stroke;  
KW gamma-aminobutyric acid receptor; allosteric modulator; antagonist;  
KW spinal cord injury; epilepsy; ischaemia; hypoglycaemia; anoxia;  
KW Alzheimer's disease; hyperparathyroidism; osteoporosis; depression;  
KW cognitive disorder.

XX Homo sapiens.

OS

XX MO9951641-A1.  
 PN 14-OCT-1999.  
 PD 02-APR-1999; 99WO-US07333.  
 PF 03-APR-1998; 98US-0080671.  
 PR (NPS-) NPS. PHARM INC.  
 PA Stormann TM, Hammerland LG, Storjohann LT, Busby JG, Garrett JE;  
 PI Slinin RT;  
 PI WPI; 1999-610995/52.  
 DR N-PSDB; AA231058.  
 DR  
 XX  
 PT New G-protein fusion receptors and chimeras containing domains from  
 PT different receptors, used to screen for modulators, potentially useful  
 PT e.g. for treating or preventing stroke or Alzheimer's disease  
 PS  
 XX Example 1; Fig 12; 255pp; English.

CC The invention relates to G-protein fusion receptors (I) comprising:  
 CC (1) in the N to C direction, extracellular (ECD), transmembrane (TMD) and  
 CC intracellular (ICD) domains, each chosen independently from a CAR  
 CC (calcium receptor), GUR (metabotropic glutamate receptor) and GABAB.  
 CC (gamma-aminobutyric acid receptor); (2) an optional linker attached to  
 CC the C-terminus of ICD; and (3) a G-protein (GP) linked to ICD or the  
 CC linker. (I), and recombinant chimeric receptors (CR) without the GP  
 CC component, are used to assess function of the various domains and to  
 CC identify compounds (e.g. allosteric modulators or antagonists) that act  
 CC on these domains. The modulators are potentially useful for treating or  
 CC preventing diseases associated with the receptors, e.g. stroke, head or  
 CC spinal cord injury, epilepsy, ischaemia, hypoglycaemia, anxiety,  
 CC Alzheimer's disease, hyperparathyroidism, osteoporosis, cognitive  
 CC disorders and depression. Nucleic acid (II) that encodes (i) is used:  
 CC (1) for recombinant production of corresponding proteins; and (2) to  
 CC produce cells used in screening for modulators. Use of CAR and mglur  
 CC domains allows presentation of GABAB domains, to a binding agent, in a  
 CC form more like the natural domain structure compared with use of  
 CC incomplete receptors, lacking one or more domains. By shuffling different  
 CC domains, agents can be identified that affect particular domains of a  
 CC receptor.

XX Sequence 1276 AA:

Query Match 98.2%; Score 1827; DB 20; Length 1276;  
 Best Local Similarity 98.6%; Pred. No. 4.3e-172;  
 Matches 353; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MTEESIMACCLSEAKERRINDEIERHVRDRKRDAREKLILLGSGSGKSTFIKMR 60  
 DB 918 MTEESIMACCLSEAKERRINDEIERHVRDRKRDAREKLILLGSGSGKSTFIKMR 977  
 QY 61 IHGSGYSDDEKRGFTKLVYQNIPTAMQAMIRAMDTLKIPYKKEHNAHAQVREVDEK 120  
 DB 978 IHGSGYSDDEKRGFTKLVYQNIPTAMQAMIRAMDTLKIPYKKEHNAHAQVREVDEK 1037  
 QY 121 VSAFENYVVAISLWMDPGIOECYDRRREYQSDSTKYYLNDLDRVADPAVLPDQDVL 180  
 DB 1038 VSAFENYVVAISLWMDPGIOECYDRRREYQSDSTKYYLNDLDRVADPAVLPDQDVL 1097  
 QY 181 RVRPPTGIEYEPDLOSIVFRWADVGGQSERKRWIHCFENYTSIMFLVALSEYQDVLV 240  
 DB 1098 RVRPPTGIEYEPDLOSIVFRWADVGGQSERKRWIHCFENYTSIMFLVALSEYQDVLV 1157  
 QY 241 ESDNENMEESKALFTITITYPWFQNSVILLFKKDLLEKIMYSHLVDFPEYDGPOR 300  
 DB 1158 ESDNENMEESKALFTITITYPWFQNSVILLFKKDLLEKIMYSHLVDFPEYDGPOR 1217  
 QY 301 DAQAREFIIKMEVDLNPDSDKITISHFTCATDTENIRVFAAVKDTIILNLKCGL 358  
 DB 1217 DAQAREFIIKMEVDLNPDSDKITISHFTCATDTENIRVFAAVKDTIILNLKCGL 358

DB 1218 DAQAREFIIKMEVDLNPDSDKITISHFTCATDTENIRVFAAVKDTIILNLKCGL 1275  
 RESULT 6  
 ID AA015093  
 ID AA015093 standard; Protein: 1276 AA.  
 AC AA015093;  
 AC 22-AUG-2002 (first entry)  
 DT  
 DE Human phCAR/hmGUR2\*Gq15 fusion construct protein.  
 DE  
 XX Human: G-protein fusion receptor; extracellular domain;  
 KW transmembrane domain; intracellular domain; CAR; mglur; GABAB;  
 KW modulator identification.  
 XX  
 OS Chimeric - Homo sapiens.  
 OS  
 XX W0200229033-A2.  
 PN W0200229033-A2.  
 PD 11-APR-2002.  
 PD  
 XX 03-OCT-2001; 2001WO-US31074.  
 PF 03-OCT-2000; 2000US-0679664.  
 PR (NPS-) NPS. PHARM INC.  
 PA Stormann T, Hammerland LG, Storjohann LT, Busby JG, Garrett JE;  
 PI Slinin RT;  
 PI WPI; 2002-330170/36.  
 DR  
 XX Novel G-protein fusion receptor, useful for identifying modulators of  
 PT CAR, mglur and GABAB, comprises G-protein joined to the intracellular  
 PT domain of the receptor -  
 XX

PS Example 3; Fig 12; 168pp; English.

CC The invention comprises G-protein fusion receptors - comprising  
 CC extracellular, transmembrane and intracellular domains similar to CAR,  
 CC mglur or GABAB receptor sequences. The G-protein fusion receptors of the  
 CC invention may also possess a linker joined to the carboxy terminus of the  
 CC intracellular domain, and a G-protein joined to the linker. The G-protein  
 CC fusion receptors of the invention are useful for identifying modulators  
 CC of CAR, mglur and GABAB for use in treating associated conditions. The  
 CC present amino acid sequence was used in the production of the invention.  
 CC  
 XX Sequence 1276 AA:

Query Match 98.2%; Score 1827; DB 23; Length 1276;  
 Best Local Similarity 98.6%; Pred. No. 4.3e-172;  
 Matches 353; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MTEESIMACCLSEAKERRINDEIERHVRDRKRDAREKLILLGSGSGKSTFIKMR 60  
 DB 918 MTEESIMACCLSEAKERRINDEIERHVRDRKRDAREKLILLGSGSGKSTFIKMR 977  
 QY 61 IHGSGYSDDEKRGFTKLVYQNIPTAMQAMIRAMDTLKIPYKKEHNAHAQVREVDEK 120  
 DB 978 IHGSGYSDDEKRGFTKLVYQNIPTAMQAMIRAMDTLKIPYKKEHNAHAQVREVDEK 1037  
 QY 121 VSAFENYVVAISLWMDPGIOECYDRRREYQSDSTKYYLNDLDRVADPAVLPDQDVL 180  
 DB 1038 VSAFENYVVAISLWMDPGIOECYDRRREYQSDSTKYYLNDLDRVADPAVLPDQDVL 1097  
 QY 181 RVRPPTGIEYEPDLOSIVFRWADVGGQSERKRWIHCFENYTSIMFLVALSEYQDVLV 240  
 DB 1098 RVRPPTGIEYEPDLOSIVFRWADVGGQSERKRWIHCFENYTSIMFLVALSEYQDVLV 1157  
 QY 241 ESDNENMEESKALFTITITYPWFQNSVILLFKKDLLEKIMYSHLVDFPEYDGPOR 300  
 DB 1217 ESDNENMEESKALFTITITYPWFQNSVILLFKKDLLEKIMYSHLVDFPEYDGPOR 300



DB 1158 ESDNENRMEESKALFRITITYPWFONSSVILFLNKKDLLEKIMYSHLVDFPEYDGPQR 1217  
QY 301 DAQAREFILKMFVDLNDPSDKIIYSHTCATDTENIRFVPAVKDTILQNLKEYNL 358  
DB 1218 DAQAREFILKMFVDLNDPSDKIIYSHTCATDTENIRFVPAVKDTILQNLKDCGL 1275

RESULT 7  
AA049129  
ID AA049129 standard; Protein: 1394 AA.  
XX  
AC AA049129;  
XX  
DT 07-JAN-2000 (first entry)  
XX  
DE pmGluR2/Car\*galphag15 fusion construct protein sequence.  
XX  
KW G-protein fusion receptor; Car; calcium receptor; GluR; head injury;  
KW metabotropic glutamate receptor; GABAR; chimeric receptor; stroke;  
KW gamma-aminobutyric acid receptor; allosteric modulator; antagonist;  
KW spinal cord injury; epilepsy; ischaemia; hypoglycaemia; anoxia;  
KW Alzheimer's disease; hyperparathyroidism; osteoporosis; depression;  
KW cognitive disorder.  
XX  
OS Homo sapiens.  
XX  
PN MO9951641-A1.  
XX  
PD 14-OCT-1999.  
XX  
PF 02-APR-1999; 99WO-US07333.  
XX  
PR 03-APR-1998; 98US-0080671.  
XX  
PA (NPS-) NPS PHARM INC.  
XX  
PI Stormann TM, Hammerland LG, Storjohann LL, Busby JG, Garrett JE;  
PI Stiml RT;  
XX  
DR WPI: 1999-610995/52.  
DR N-PSDB: AA231060.  
XX  
XX New G-protein fusion receptors and chimeras containing domains from  
PT different receptors, used to screen for modulators, potentially useful  
PT e.g. for treating or preventing stroke or Alzheimer's disease -  
XX  
PS Example 1: Fig 12; 255pp; English.

The invention relates to G-protein fusion receptors (I) comprising:  
(1) in the N to C direction, extracellular (ECD), transmembrane (TMD) and  
intracellular (ICD) domains, each chosen independently from a Car  
(calcium receptor), GluR (metabotropic glutamate receptor) and GABAR  
(gamma-aminobutyric acid receptor); (2) an optional linker attached to  
the C-terminus of ICD; and (3) a G-protein (GP) linked to ICD or the  
linker. (1), and recombinant chimeric receptors (CR) without the GP  
component, are used to assess function of the various domains and to  
identify compounds (e.g. allosteric modulators or antagonists) that act  
on these domains. The modulators are potentially useful for treating or  
preventing diseases associated with the receptors, e.g. stroke, head or  
spinal cord injury, epilepsy, ischaemia, hypoglycaemia, anoxia,  
Alzheimer's disease, hyperparathyroidism, osteoporosis, cognitive  
disorders and depression. Nucleic acid (II) that encodes (I) is used:  
(1) for recombinant production of corresponding proteins; and (2) to  
produce cells used in screening for modulators. Use of Car and mGluR  
domains allows presentation of GABAR domains, to a binding agent, in a  
form more like the natural domain structure compared with use of  
incomplete receptors, lacking one or more domains. By shuffling different  
domains, agents can be identified that affect particular domains of a  
receptor.

Sequence 1394 AA;  
Query Match 98.2%; Score 1827; DB 20; Length 1394;

Best Local Similarity 98.6%; Pred. No. 4.9e-172;  
Matches 353; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
QY 1 MTLSEIMACCSSEAKKARRINDEIERHVRDKDARRELKLLLGSGSKSTFIQMR 60  
DB 1036 MTLSEIMACCSSEAKKARRINDEIERDLRRDKDARRELKLLLGSGSKSTFIQMR 1095  
QY 61 ITHSGYSDEDKRGFTKLVYONIFPAMQAMIRAMDILKIPKYEHKNAHQLVREVDYK 120  
DB 1096 ITHSGYSDEDKRGFTKLVYONIFPAMQAMIRAMDILKIPKYEHKNAHQLVREVDYK 1155  
QY 121 VSAFENPYVDARKSLMNDPGIOECYDRRREYQLSDSTKYLYLNDLDRVADPAVLPDQDVL 180  
DB 1156 VSAFENPYVDARKSLMNDPGIOECYDRRREYQLSDSTKYLYLNDLDRVADPAVLPDQDVL 1215  
QY 181 RVRVPTTGIIYPPDLOSIVIRFMDVGGQSRERKWHICFENVTSTMFVALSEYDQVLV 240  
DB 1216 RVRVPTTGIIYPPDLOSIVIRFMDVGGQSRERKWHICFENVTSTMFVALSEYDQVLV 1275  
QY 241 ESDNENRMEESKALFRITITYPWFONSSVILFLNKKDLLEKIMYSHLVDFPEYDGPQR 300  
DB 1276 ESDNENRMEESKALFRITITYPWFONSSVILFLNKKDLLEKIMYSHLVDFPEYDGPQR 1335  
QY 301 DAQAREFILKMFVDLNDPSDKIIYSHTCATDTENIRFVPAVKDTILQNLKEYNL 358  
DB 1336 DAQAREFILKMFVDLNDPSDKIIYSHTCATDTENIRFVPAVKDTILQNLKDCGL 1393

RESULT 8  
AA015095  
ID AA015095 standard; Protein: 1394 AA.  
XX  
AC AA015095;  
XX  
DT 22-AUG-2002 (first entry)  
XX  
DE Human pmGluR2-Car\*G-alpha-q15 fusion construct protein.  
XX  
KW Human; G-protein fusion receptor; extracellular domain;  
KW transmembrane domain; intracellular domain; Car; mGluR; GABAR;  
KW modulator identification.  
XX  
OS Chimeric - Homo sapiens.  
XX  
PN MO200229033-A2.  
XX  
PD 11-APR-2002.  
XX  
PF 03-OCT-2001; 2001WO-US31074.  
XX  
PR 03-OCT-2000; 2000US-0679664.  
XX  
PA (NPS-) NPS PHARM INC.  
XX  
PI Stormann T, Hammerland LG, Storjohann LL, Busby JG, Garrett JE;  
PI Stiml RT;  
XX  
DR WPI: 2002-330170/36.

Novel G-protein fusion receptor, useful for identifying modulators of  
Car, mGluR and GABAR, comprises G-protein joined to the intracellular  
domain of the receptor -  
XX  
PS Example 3: Fig 12; 168pp; English.

The invention comprises G-protein fusion receptors - comprising  
extracellular, transmembrane and intracellular domains similar to Car,  
mGluR or GABAR receptor sequences. The G-protein fusion receptors of the  
invention may also possess a linker joined to the carboxy terminus of the  
intracellular domain, and a G-protein joined to the linker. The G-protein  
fusion receptors of the invention are useful for identifying modulators  
of Car, mGluR and GABAR for use in treating associated conditions. The  
present amino acid sequence was used in the production of the invention.

```

XX Sequence      1394 AA:
SQ
Query Match      98.2%; Score 1827; DB 23; Length 1394;
Best Local Similarity 98.6%; Pred. No. 4.9e-172;
Matches 353; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MTEISIMACCLSEAKKARRINDEIERHVRDKRDARRELKLLLGTEGSGKSTFIKQMR 60
DB MTEISIMACCLSEAKKARRINDEIERHVRDKRDARRELKLLLGTEGSGKSTFIKQMR 1095
QY 61 ITHSGSYDEDKRGFTKLVYQNIPTAMQAMIRAMDTLKIPYKYEHNKAHAQVREVDYER 120
DB ITHSGSYDEDKRGFTKLVYQNIPTAMQAMIRAMDTLKIPYKYEHNKAHAQVREVDYER 1155
QY 121 VSAFENPYDAIKSLMDPGIOECYDRRREYQSDSKTYLNDLRVADPAVLPYDQDVL 180
DB 1156 VSAFENPYDAIKSLMDPGIOECYDRRREYQSDSKTYLNDLRVADPAVLPYDQDVL 1215
QY 181 RVRVPTTGIIIEYPPDLQSVIFRMYDVGQSRERKRWIHCFENVTSMFLVALSEYDQVLY 240
DB 1216 RVRVPTTGIIIEYPPDLQSVIFRMYDVGQSRERKRWIHCFENVTSMFLVALSEYDQVLY 1275
QY 241 ESDNENMEESKALFRITITYPWFQNSVILFLNKKDLLEKIMYSHLVDPPEYDGPOR 300
DB 1276 ESDNENMEESKALFRITITYPWFQNSVILFLNKKDLLEKIMYSHLVDPPEYDGPOR 1335
QY 301 DAQARFEILKMFVDLNPDSDKIITYSHFTCATDTENTIRFVFAAVKDTIQLNLKEYNL 358
DB 1336 DAQARFEILKMFVDLNPDSDKIITYSHFTCATDTENTIRFVFAAVKDTIQLNLKEYNL 1393

RESULT 9
AAV49134
ID AAV49134 standard; Protein: 1397 AA.
XX
AC AAV49134:
XX
DT 07-JAN-2000 (first entry)
XX
DE pmg1ur2/Gar*Gal*alpha15+3A1a linker fusion construct protein sequence.
XX
KW G-protein fusion receptor; CAR; calcium receptor; G1UR; head injury;
KW metabotropic glutamate receptor; GABABR; chimeric receptor; stroke;
KW gamma-aminobutyric acid receptor; allosteric modulator; antagonist;
KW spinal cord injury; epilepsy; ischaemia; hypoglycaemia; anoxia;
KW Alzheimer's disease; hyperparathyroidism; osteoporosis; depression;
KW cognitive disorder.
XX
OS Homo sapiens.
XX
PN WO9951641-A1.
XX
PD 14-OCT-1999.
XX
PF 02-APR-1999; 99WO-US07333.
XX
PR 03-APR-1998; 98US-0080671.
XX
PA (NPSF-) NPS PHARM INC.
XX
PI Stormann TM, Hammerland LG, Storchmann LL, Busby JG, Garrett JE;
PI Stimin RT;
XX
DR WPI, 1999-610995/52.
XX
DR N-PSDB; AA231065.
XX
PT New G-protein fusion receptors and chimeras containing domains from
XX different receptors, used to screen for modulators, potentially useful
XX e.g. for treating or preventing stroke or Alzheimer's disease
XX
XX Example 1; Fig 12; 255pp; English.
XX

```

```

CC The invention relates to G-protein fusion receptors (I) comprising:
CC (1) in the N to C direction, extracellular (ECD), transmembrane (TMD) and
CC intracellular (ICD) domains, each chosen independently from a Car
CC (calcium receptor), G1UR (metabotropic glutamate receptor) and GABABR
CC (gamma-aminobutyric acid receptor); (2) an optional linker attached to
CC the C-terminus of ICD; and (3) a G-protein (GP) linked to ICD or the
CC linker; (1), and recombinant chimeric receptors (CR) without the GP
CC component, are used to assess function of the various domains and to
CC identify compounds (e.g. allosteric modulators or antagonists) that act
CC on these domains. The modulators are potentially useful for treating or
CC preventing diseases associated with the receptors, e.g. stroke, head or
CC spinal cord injury, epilepsy, ischaemia, hypoglycaemia, anoxia,
CC Alzheimer's disease, hyperparathyroidism, osteoporosis, cognitive
CC disorders and depression. Nucleic acid (II) that encodes (1) is used:
CC (1) for recombinant production of corresponding proteins; and (2) to
CC produce cells used in screening for modulators. Use of Car and mg1ur
CC domains allows presentation of GABABR domains, to a binding agent. In a
CC form more like the natural domain structure compared with use of
CC incomplete receptors, lacking one or more domains. By shuffling different
CC domains, agents can be identified that affect particular domains of a
CC receptor.
CC
SQ Sequence      1397 AA:
Query Match      98.2%; Score 1827; DB 20; Length 1397;
Best Local Similarity 98.6%; Pred. No. 4.9e-172;
Matches 353; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MTEISIMACCLSEAKKARRINDEIERHVRDKRDARRELKLLLGTEGSGKSTFIKQMR 60
DB MTEISIMACCLSEAKKARRINDEIERHVRDKRDARRELKLLLGTEGSGKSTFIKQMR 1098
QY 61 ITHSGSYDEDKRGFTKLVYQNIPTAMQAMIRAMDTLKIPYKYEHNKAHAQVREVDYER 120
DB ITHSGSYDEDKRGFTKLVYQNIPTAMQAMIRAMDTLKIPYKYEHNKAHAQVREVDYER 1158
QY 121 VSAFENPYDAIKSLMDPGIOECYDRRREYQSDSKTYLNDLRVADPAVLPYDQDVL 180
DB 1159 VSAFENPYDAIKSLMDPGIOECYDRRREYQSDSKTYLNDLRVADPAVLPYDQDVL 1218
QY 181 RVRVPTTGIIIEYPPDLQSVIFRMYDVGQSRERKRWIHCFENVTSMFLVALSEYDQVLY 240
DB 1219 RVRVPTTGIIIEYPPDLQSVIFRMYDVGQSRERKRWIHCFENVTSMFLVALSEYDQVLY 1278
QY 241 ESDNENMEESKALFRITITYPWFQNSVILFLNKKDLLEKIMYSHLVDPPEYDGPOR 300
DB 1279 ESDNENMEESKALFRITITYPWFQNSVILFLNKKDLLEKIMYSHLVDPPEYDGPOR 1338
QY 301 DAQARFEILKMFVDLNPDSDKIITYSHFTCATDTENTIRFVFAAVKDTIQLNLKEYNL 358
DB 1339 DAQARFEILKMFVDLNPDSDKIITYSHFTCATDTENTIRFVFAAVKDTIQLNLKEYNL 1396

RESULT 10
AAO15100
ID AAO15100 standard; Protein: 1397 AA.
XX
AC AAO15100:
XX
DT 22-AUG-2002 (first entry)
XX
DE Human pmg1ur2-Car*G-alpha-q15+3A1a linker fusion construct protein.
XX
KW Human; G-protein fusion receptor; extracellular domain;
KW transmembrane domain; intracellular domain; CAR; mg1ur; GABABR;
KW modulator identification.
XX
OS Chimeric - Homo sapiens.
XX
OS Chimeric - unidentified.
XX
PN WO200229033-A2.
XX
PD 11-APR-2002.
XX

```

XX 03-OCT-2001; 2001WO-US31074.  
PF  
XX  
XX 03-OCT-2000; 2000US-0679664.  
PR  
XX  
XX (NPSF-) NPS PHARM INC.  
PA  
PI Stormann T, Hammerland LG, Storjohann LL, Busby JG, Garrett JE;  
PI Simin RT;  
XX WPI; 2002-330170/36.  
DR  
XX  
XX Novel G-protein fusion receptor, useful for identifying modulators of  
PT Car, mglur and GABAB, comprises G-protein joined to the intracellular  
PT domain of the receptor -  
XX  
XX Disclosure: Fig 12; 168pp; English.  
PS  
XX  
XX The invention comprises G-protein fusion receptors - comprising  
CC extracellular, transmembrane and intracellular domains similar to Car,  
CC mglur or GABAB receptor sequences. The G-protein fusion receptors of the  
CC invention may also possess a linker joined to the carboxy terminus of the  
CC intracellular domain, and a G-protein joined to the linker. The G-protein  
CC fusion receptors of the invention are useful for identifying modulators  
CC of Car, mglur and GABAB for use in treating associated conditions. The  
CC present amino acid sequence was used in the production of the invention.  
SQ Sequence 1397 AA;

Query Match 98.2%; Score 1827; DB 23; Length 1397;  
Best Local Similarity 98.6%; Pred. No. 4,9e-172;  
Matches 353; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MLESIMACCLSEEAKEARRINDEIERHVRDRKRDARELKLILGTGSGKSTFIKQMR 60  
DB 1039 MLESIMACCLSEEAKEARRINDEIERHVRDRKRDARELKLILGTGSGKSTFIKQMR 1098  
QY 61 IIHSGSYDEDEKRGFTKLIVQNIPTAMQAMIRAMDTLKIPIKYEHNKKAHAQLVREVDYK 120  
DB 1099 IIHSGSYDEDEKRGFTKLIVQNIPTAMQAMIRAMDTLKIPIKYEHNKKAHAQLVREVDYK 1158  
QY 121 VSAFENPYVDAKSLMNDPGIOECYDRREYQLSDSTKYLLNDLDRVADPAVLPTQOQVL 180  
DB 1159 VSAFENPYVDAKSLMNDPGIOECYDRREYQLSDSTKYLLNDLDRVADPAVLPTQOQVL 1218  
QY 181 RVRVPTGTGIIIEYFPDQSVIFPMVDVGORSERKWKHCFENVTSIMFLVALSEYDQVL 240  
DB 1219 RVRVPTGTGIIIEYFPDQSVIFPMVDVGORSERKWKHCFENVTSIMFLVALSEYDQVL 1278  
QY 241 ESDNENRMEESKALFRITITTPWFQNSSVILFLNKKDLLEEKIMYSHLVDFPEYDGPQR 300  
DB 1279 ESDNENRMEESKALFRITITTPWFQNSSVILFLNKKDLLEEKIMYSHLVDFPEYDGPQR 1338  
QY 301 DAQAAREFTLKMFLVNLNPDSDKIITSHFTCATDTENIRFVPAAYVDTIQLNLKQYNL 358  
DB 1339 DAQAAREFTLKMFLVNLNPDSDKIITSHFTCATDTENIRFVPAAYVDTIQLNLKQYNL 1396

RESULT 11  
AA015105  
ID AA015105 standard; Protein; 1402 AA.

XX AA015105;

XX 22-AUG-2002 (first entry)

XX Human ph2SPMglur3-Car\*AAA\*Gq15 fusion construct protein sequence.

XX Human; G-protein fusion receptor; extracellular domain;

XX transmembrane domain; intracellular domain; Car; mglur; GABAB;

XX modulator identification.

XX Chimeric - Homo sapiens.

XX WO200229033-A2.  
PN  
XX  
XX 11-APR-2002.  
PD  
XX  
XX 03-OCT-2001; 2001WO-US31074.  
PF  
XX  
XX 03-OCT-2000; 2000US-0679664.  
PR  
XX  
XX (NPSF-) NPS PHARM INC.  
PA  
PI Stormann T, Hammerland LG, Storjohann LL, Busby JG, Garrett JE;  
PI Simin RT;  
XX WPI; 2002-330170/36.  
DR N-PSDB; AAL43286.  
XX  
XX  
XX Novel G-protein fusion receptor, useful for identifying modulators of  
PT Car, mglur and GABAB, comprises G-protein joined to the intracellular  
PT domain of the receptor -  
XX  
XX Disclosure: Fig 18; 168pp; English.  
PS  
XX  
XX The invention comprises G-protein fusion receptors - comprising  
CC extracellular, transmembrane and intracellular domains similar to Car,  
CC mglur or GABAB receptor sequences. The G-protein fusion receptors of the  
CC invention may also possess a linker joined to the carboxy terminus of the  
CC intracellular domain, and a G-protein joined to the linker. The G-protein  
CC fusion receptors of the invention are useful for identifying modulators  
CC of Car, mglur and GABAB for use in treating associated conditions. The  
CC present amino acid sequence was used in the production of the invention.  
SQ Sequence 1402 AA;

Query Match 98.2%; Score 1827; DB 23; Length 1402;  
Best Local Similarity 98.6%; Pred. No. 4,9e-172;  
Matches 353; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MLESIMACCLSEEAKEARRINDEIERHVRDRKRDARELKLILGTGSGKSTFIKQMR 60  
DB 1044 MLESIMACCLSEEAKEARRINDEIERHVRDRKRDARELKLILGTGSGKSTFIKQMR 1103  
QY 61 IIHSGSYDEDEKRGFTKLIVQNIPTAMQAMIRAMDTLKIPIKYEHNKKAHAQLVREVDYK 120  
DB 1104 IIHSGSYDEDEKRGFTKLIVQNIPTAMQAMIRAMDTLKIPIKYEHNKKAHAQLVREVDYK 1163  
QY 121 VSAFENPYVDAKSLMNDPGIOECYDRREYQLSDSTKYLLNDLDRVADPAVLPTQOQVL 180  
DB 1164 VSAFENPYVDAKSLMNDPGIOECYDRREYQLSDSTKYLLNDLDRVADPAVLPTQOQVL 1223  
QY 181 RVRVPTGTGIIIEYFPDQSVIFPMVDVGORSERKWKHCFENVTSIMFLVALSEYDQVL 240  
DB 1224 RVRVPTGTGIIIEYFPDQSVIFPMVDVGORSERKWKHCFENVTSIMFLVALSEYDQVL 1283  
QY 241 ESDNENRMEESKALFRITITTPWFQNSSVILFLNKKDLLEEKIMYSHLVDFPEYDGPQR 300  
DB 1284 ESDNENRMEESKALFRITITTPWFQNSSVILFLNKKDLLEEKIMYSHLVDFPEYDGPQR 1343  
QY 301 DAQAAREFTLKMFLVNLNPDSDKIITSHFTCATDTENIRFVPAAYVDTIQLNLKQYNL 358  
DB 1344 DAQAAREFTLKMFLVNLNPDSDKIITSHFTCATDTENIRFVPAAYVDTIQLNLKQYNL 1401

RESULT 12

XX AA49131 standard; Protein; 1418 AA.

XX AA49131;

XX 07-JAN-2000 (first entry)

XX mglur8/Car\*Galphag15 fusion construct protein sequence.

XX

KW G-protein fusion receptor; Car; calcium receptor; Glur; head injury;  
 KW metabotropic glutamate receptor; GABAB; chimeric receptor; stroke;  
 KW gamma-aminobutyric acid receptor; allosteric modulator; antagonist;  
 KW spinal cord injury; epilepsy; ischaemia; hypoglycaemia; anoxia;  
 KW Alzheimer's disease; hyperparathyroidism; osteoporosis; depression;  
 KW cognitive disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN MO9951641-A1.  
 XX  
 PD 14-OCT-1999.  
 XX  
 PF 02-APR-1999; 99WO-US07333.  
 XX  
 PR 03-APR-1998; 98US-0080671.  
 XX  
 PA (NPS- ) NPS PHARM INC.  
 XX  
 PI Stormann TM, Hammerland LG, Storjohann LL, Busby JG, Garrett JE;  
 PI Stimin RT;  
 DR WPI: 1999-610995/52.  
 DR N-PSDB: AAZ31062.  
 XX  
 PT New G-protein fusion receptors and chimeras containing domains from  
 PT different receptors, used to screen for modulators, potentially useful  
 PT e.g. for treating or preventing stroke or Alzheimer's disease  
 XX  
 PS Example 1; Fig 12; 255pp; English.  
 XX  
 CC The invention relates to G-protein fusion receptors (I) comprising:  
 CC (1) in the N to C direction, extracellular (ECD), transmembrane (TMD) and  
 CC intracellular (ICD) domains, each chosen independently from a Car  
 CC (calcium receptor), Glur (metabotropic glutamate receptor) and GABAB  
 CC (gamma-aminobutyric acid receptor); (2) an optional linker attached to  
 CC the C-terminus of ICD; and (3) a G-protein (GP) linked to ICD or the  
 CC linker. (I), and recombinant chimeric receptors (CR) without the GP  
 CC component, are used to assess function of the various domains and to  
 CC identify compounds (e.g. allosteric modulators or antagonists) that act  
 CC on these domains. The modulators are potentially useful for treating or  
 CC preventing diseases associated with the receptors, e.g. stroke, head or  
 CC spinal cord injury, epilepsy, ischaemia, hypoglycaemia, anoxia,  
 CC Alzheimer's disease, hyperparathyroidism, osteoporosis, cognitive  
 CC disorders and depression. Nucleic acid (II) that encodes (I) is used:  
 CC (1) for recombinant production of corresponding proteins; and (2) to  
 CC produce cells used in screening for modulators. Use of Car and mglur  
 CC domains allows presentation of GABAB domains, to a binding agent, in a  
 CC form more like the natural domain structure compared with use of  
 CC incomplete receptors, lacking one or more domains. By shuffling different  
 CC domains, agents can be identified that affect particular domains of a  
 CC receptor.  
 XX  
 XX Sequence 1418 AA:  
 SQ  
 Query Match 98.2%; Score 1827; DB 20; Length 1418;  
 Best Local Similarity 98.6%; Pred. No. 5e-172;  
 Matches 353; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

DB 1240 RVRVPTTGIIIEYPPDLSVIFRWVDVGQNSERRKMHCFBNVTSIMEFLVALSEYDVLV 1299  
 QY 241 ESDNENRNEESKALFRITITYPWFQNSVILFLNKKDLLEKIMYSHLVDFPEYDGPOR 300  
 DB 1300 ESDNENRNEESKALFRITITYPWFQNSVILFLNKKDLLEKIMYSHLVDFPEYDGPOR 1359  
 QY 301 DAQAREFIILKMFVLDLPDSDSKITISHTCTDPTENIRFVPAANKDTILNLKEYNL 358  
 DB 1360 DAQAREFIILKMFVLDLPDSDSKITISHTCTDPTENIRFVPAANKDTILNLKDCGL 1417  
 RESULT 13  
 ID AA015097 standard; Protein: 1418 AA.  
 XX  
 AC AA015097;  
 XX  
 DT 22-AUG-2002 (first entry)  
 XX  
 DE Human mglur8-Car+G-alpha-q15 fusion construct protein.  
 XX  
 KW Human; G-protein fusion receptor; extracellular domain;  
 KW transmembrane domain; intracellular domain; Car; mglur; GABAB;  
 KW modulator identification.  
 XX  
 OS Chimeric - Homo sapiens.  
 XX  
 PN WO200229033-A2.  
 XX  
 PD 11-APR-2002.  
 XX  
 PF 03-OCT-2001; 2001WO-US31074.  
 XX  
 PR 03-OCT-2000; 2000US-0679664.  
 XX  
 PA (NPS- ) NPS PHARM INC.  
 XX  
 PI Stormann T, Hammerland LG, Storjohann LL, Busby JG, Garrett JE;  
 PI Stimin RT;  
 DR WPI: 2002-330170/36.  
 DR  
 XX  
 PT Novel G-protein fusion receptor, useful for identifying modulators of  
 PT Car, mglur and GABAB, comprises G-protein joined to the intracellular  
 PT domain of the receptor -  
 XX  
 PS Example 3; Fig 12; 168pp; English.  
 XX  
 CC The invention comprises G-protein fusion receptors - comprising  
 CC extracellular, transmembrane and intracellular domains similar to Car,  
 CC mglur or GABAB receptor sequences. The G-protein fusion receptors of the  
 CC invention may also possess a linker joined to the carboxy terminus of the  
 CC intracellular domain, and a G-protein joined to the linker. The G-protein  
 CC fusion receptors of the invention are useful for identifying modulators  
 CC of Car, mglur and GABAB for use in treating associated conditions. The  
 CC present amino acid sequence was used in the production of the invention.  
 XX  
 XX Sequence 1418 AA:  
 SQ  
 Query Match 98.2%; Score 1827; DB 23; Length 1418;  
 Best Local Similarity 98.6%; Pred. No. 5e-172;  
 Matches 353; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Db 1180 VSAFENPVDAIKSLMNDPGIOECYDRRREYQLSDSTKYLYLNDLDRVADPAVLPFOQDVL 1239  
QY 181 RVRVPTGIIIEYPPDLQSVIFRMVDVGQSRERKWHCFENWTSIMFLVALSEYDQVL 240  
Db 1240 RVRVPTGIIIEYPPDLQSVIFRMVDVGQSRERKWHCFENWTSIMFLVALSEYDQVL 1239  
QY 241 ESDNENRMEESKALFRITITYPWFONSSVILFLNKKDLLEKIMYSHLVDFPEYDGPQR 300  
Db 1300 ESDNENRMEESKALFRITITYPWFONSSVILFLNKKDLLEKIMYSHLVDFPEYDGPQR 1359  
QY 301 DAQAREFILKMFVDLNPDSKIIYSHTCATDTENIRFVFAAVKDTIQLNLKCYNL 358  
Db 1360 DAQAREFILKMFVDLNPDSKIIYSHTCATDTENIRFVFAAVKDTIQLNLKDCGL 1417

RESULT 14  
AA015103  
ID AA015103 standard; Protein: 1421 AA.  
XX AA015103;  
XX  
XX  
XX 22-AUG-2002 (first entry)  
XX  
XX  
XX Human phmGluR8-Car\*AA\*G-alpha-q15 fusion protein.  
DE  
XX Human; G-protein fusion receptor; extracellular domain;  
KW transmembrane domain; intracellular domain; Car; mGluR; GABAB;  
KW modulator identification.  
XX  
XX Chimeric - Homo sapiens.  
OS  
XX WO200229033-A2.  
XX  
XX 11-APR-2002.  
XX  
XX 03-OCT-2001; 2001WO-US31074.  
XX  
XX 03-OCT-2000; 2000US-0679664.  
XX  
XX  
XX (NPS-) NPS PHARM INC.  
PA  
XX Stormann T, Hammerland LG, Storjohann LL, Busby JG, Garrett JE;  
PI Simin RT;  
XX  
XX WPI: 2002-330170/36.  
XX  
XX  
XX Novel G-protein fusion receptor, useful for identifying modulators of  
PT Car, mGluR and GABAB, comprises G-protein joined to the intracellular  
PT domain of the receptor -  
XX  
XX  
XX PS Disclosure: Fig 16; 168pp; English.  
XX  
XX The invention comprises G-protein fusion receptors - comprising  
CC extracellular, transmembrane and intracellular domains similar to Car,  
CC mGluR or GABAB receptor sequences. The G-protein fusion receptors of the  
CC invention may also possess a linker joined to the carboxy terminus of the  
CC intracellular domain, and a G-protein joined to the linker. The G-protein  
CC fusion receptors of the invention are useful for identifying modulators  
CC of Car, mGluR and GABAB for use in treating associated conditions. The  
CC present amino acid sequence was used in the production of the invention.  
XX  
XX  
XX SO Sequence 1421 AA;  
Query Match 98.2%; Score 1827; DB 23; Length 1421;  
Best Local Similarity 98.6%; Pred. No. 5e-172;  
Matches 353; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MLESIMACCLSEAKARRINDEIRHVRDKRDARELKLILGSGSGSTFIKQMR 60  
Db 1063 MLESIMACCLSEAKARRINDEIRHVRDKRDARELKLILGSGSGSTFIKQMR 1122  
QY 61 IHHGGYDEDEKRGFTKLYONIFTAQAMIRAMDTLKIPYKEHNKAHQLVREVDYER 120  
IIHHGGYDEDEKRGFTKLYONIFTAQAMIRAMDTLKIPYKEHNKAHQLVREVDYER 120

Db 1123 IHHGGYDEDEKRGFTKLYONIFTAQAMIRAMDTLKIPYKEHNKAHQLVREVDYER 1182  
QY 121 VSAFENPVDAIKSLMNDPGIOECYDRRREYQLSDSTKYLYLNDLDRVADPAVLPFOQDVL 180  
Db 1183 VSAFENPVDAIKSLMNDPGIOECYDRRREYQLSDSTKYLYLNDLDRVADPAVLPFOQDVL 1242  
QY 181 RVRVPTGIIIEYPPDLQSVIFRMVDVGQSRERKWHCFENWTSIMFLVALSEYDQVL 240  
Db 1243 RVRVPTGIIIEYPPDLQSVIFRMVDVGQSRERKWHCFENWTSIMFLVALSEYDQVL 1302  
QY 241 ESDNENRMEESKALFRITITYPWFONSSVILFLNKKDLLEKIMYSHLVDFPEYDGPQR 300  
Db 1303 ESDNENRMEESKALFRITITYPWFONSSVILFLNKKDLLEKIMYSHLVDFPEYDGPQR 1362  
QY 301 DAQAREFILKMFVDLNPDSKIIYSHTCATDTENIRFVFAAVKDTIQLNLKCYNL 358  
Db 1363 DAQAREFILKMFVDLNPDSKIIYSHTCATDTENIRFVFAAVKDTIQLNLKDCGL 1420

RESULT 15  
AA015102  
ID AA015102 standard; Protein: 1422 AA.  
XX AA015102;  
XX  
XX  
XX 22-AUG-2002 (first entry)  
XX  
XX  
XX Human phmGluR4-Car\*AA\*G-alpha-q15 fusion protein.  
DE  
XX Human; G-protein fusion receptor; extracellular domain;  
KW transmembrane domain; intracellular domain; Car; mGluR; GABAB;  
KW modulator identification.  
XX  
XX Chimeric - Homo sapiens.  
OS  
XX WO200229033-A2.  
XX  
XX 11-APR-2002.  
XX  
XX 03-OCT-2001; 2001WO-US31074.  
XX  
XX 03-OCT-2000; 2000US-0679664.  
XX  
XX  
XX (NPS-) NPS PHARM INC.  
PA  
XX Stormann T, Hammerland LG, Storjohann LL, Busby JG, Garrett JE;  
PI Simin RT;  
XX  
XX WPI: 2002-330170/36.  
XX  
XX  
XX Novel G-protein fusion receptor, useful for identifying modulators of  
PT Car, mGluR and GABAB, comprises G-protein joined to the intracellular  
PT domain of the receptor -  
XX  
XX  
XX PS Disclosure: Fig 16; 168pp; English.  
XX  
XX The invention comprises G-protein fusion receptors - comprising  
CC extracellular, transmembrane and intracellular domains similar to Car,  
CC mGluR or GABAB receptor sequences. The G-protein fusion receptors of the  
CC invention may also possess a linker joined to the carboxy terminus of the  
CC intracellular domain, and a G-protein joined to the linker. The G-protein  
CC fusion receptors of the invention are useful for identifying modulators  
CC of Car, mGluR and GABAB for use in treating associated conditions. The  
CC present amino acid sequence was used in the production of the invention.  
XX  
XX  
XX SO Sequence 1422 AA;  
Query Match 98.2%; Score 1827; DB 23; Length 1422;  
Best Local Similarity 98.6%; Pred. No. 5.1e-172;  
Matches 353; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MLESIMACCLSEAKARRINDEIRHVRDKRDARELKLILGSGSGSTFIKQMR 60  
IIHHGGYDEDEKRGFTKLYONIFTAQAMIRAMDTLKIPYKEHNKAHQLVREVDYER 120

```
Db 1064 MLESIMACCLSEAKKEARRINDEIERQLRBDKRDARRELKLLLGTSBGSKSTFIKQMR 1123
Qy 61 IIHSGYSDEDKRGFTKLVOQNIPTAMQAMIRAMDTLKIPYKEHNKAHAQLVREVDYER 120
Db 1124 IIHSGYSDEDEKRGFTKLVOQNIPTAMQAMIRAMDTLKIPYKEHNKAHAQLVREVDYER 1183
Qy 121 VSAFENPYVDAIKSINNDPGIOECYDRRREYQLSSTKYIINDLDRAVDPAIPLPTQODVL 180
Db 1184 VSAFENPYVDAIKSLMNDPGIOECYDRRREYQLSSTKYIINDLDRAVDPAIPLPTQODVL 1243
Qy 181 RVRVPTTGIIIEYPEDLOSVIFRMVDVGQSRERRKIHCFENVTSIMFLVALSEYDOVL 240
Db 1244 RVRVPTTGIIIEYPEDLOSVIFRMVDVGQSRERRKIHCFENVTSIMFLVALSEYDOVL 1303
Qy 241 ESDNENRMEESKALFRITITYPWFONSSVILFLNKKDLLEEKIMYSHLVDFPEYDGPQR 300
Db 1304 ESDNENRMEESKALFRITITYPWFONSSVILFLNKKDLLEEKIMYSHLVDFPEYDGPQR 1363
Qy 301 DAQAREFIIKMFVDLNDPDSKIIYSHFTCATDPENIRFVFAAVKDTLLQNLKEYNL 358
Db 1364 DAQAREFIIKMFVDLNDPDSKIIYSHFTCATDPENIRFVFAAVKDTLLQNLKDCGL 1421
```

Search completed: June 30, 2003, 16:16:16  
Job time : 45 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2003, 16:15:31 ; Search time 15 seconds  
(Without alignments)  
704.189 Million cell updates/sec

Title: US-09-899-295-2  
Perfect score: 1860  
Sequence: 1 MTEISIMACCLSEAKREARR.....VEAAVKDTILQLNKEYNLV 359

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

- 1: Issued\_Patents\_AA.\*
- 2: /cgn2\_6/ptodata/1/1aa/5A-COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/1aa/5B-COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/1aa/6A-COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/1aa/6B-COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/1aa/6C-COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1066	57.3	374	3 US-08-878-801-4	Sequence 4, Appli
2	1063	57.2	374	4 US-09-442-349A-29	Sequence 29, Appli
3	1063	57.2	374	4 US-09-442-349A-31	Sequence 31, Appli
4	1062	57.1	374	4 US-09-442-349A-21	Sequence 21, Appli
5	1062	57.1	374	4 US-09-442-349A-23	Sequence 23, Appli
6	1062	57.1	374	4 US-09-442-349A-25	Sequence 25, Appli
7	1062	57.1	374	4 US-09-442-349A-30	Sequence 30, Appli
8	1062	57.1	374	4 US-09-442-349A-32	Sequence 32, Appli
9	1062	57.1	374	4 US-09-442-349A-34	Sequence 34, Appli
10	1061	57.0	374	4 US-09-442-349A-22	Sequence 22, Appli
11	1061	57.0	374	4 US-09-442-349A-27	Sequence 27, Appli
12	1061	57.0	374	4 US-09-442-349A-10	Sequence 10, Appli
13	1061	57.0	374	4 US-09-442-349A-33	Sequence 33, Appli
14	1058	56.9	374	4 US-09-442-349A-26	Sequence 26, Appli
15	1058	56.9	374	4 US-09-442-349A-28	Sequence 28, Appli
16	1057	56.8	374	4 US-09-442-349A-20	Sequence 20, Appli
17	1057	56.8	374	4 US-09-442-349A-27	Sequence 27, Appli
18	1055	56.7	374	4 US-09-218-489-2	Sequence 2, Appli
19	1054	56.7	374	4 US-09-442-349A-14	Sequence 14, Appli
20	1054	56.7	374	4 US-09-442-349A-16	Sequence 16, Appli
21	1053	56.6	365	4 US-09-442-349A-108	Sequence 108, App
22	1053	56.6	374	4 US-09-442-349A-6	Sequence 6, Appli
23	1053	56.6	374	4 US-09-442-349A-8	Sequence 8, Appli
24	1053	56.6	374	4 US-09-442-349A-15	Sequence 15, Appli
25	1053	56.6	374	4 US-09-442-349A-17	Sequence 17, Appli
26	1053	56.6	374	4 US-09-442-349A-19	Sequence 19, Appli
27	1053	56.6	374	4 US-09-442-349A-19	Sequence 19, Appli

28	1052	56.6	374	4 US-09-442-349A-1	Sequence 1, Appli
29	1052	56.6	374	4 US-09-442-349A-7	Sequence 7, Appli
30	1052	56.6	374	4 US-09-442-349A-9	Sequence 9, Appli
31	1052	56.6	374	4 US-09-442-349A-18	Sequence 18, Appli
32	1049	56.4	374	4 US-09-442-349A-11	Sequence 11, Appli
33	1049	56.4	374	4 US-09-442-349A-5	Sequence 5, Appli
34	1048	56.3	374	4 US-09-442-349A-13	Sequence 13, Appli
35	1048	56.3	374	4 US-09-442-349A-12	Sequence 12, Appli
36	1046	56.2	374	4 US-09-442-349A-36	Sequence 36, Appli
37	1046	56.2	374	4 US-09-442-349A-38	Sequence 38, Appli
38	1045	56.2	374	4 US-09-442-349A-3	Sequence 3, Appli
39	1045	56.2	374	4 US-09-442-349A-37	Sequence 37, Appli
40	1045	56.2	374	4 US-09-442-349A-48	Sequence 48, Appli
41	1045	56.2	374	4 US-09-442-349A-50	Sequence 50, Appli
42	1045	56.2	374	4 US-09-442-349A-54	Sequence 54, Appli
43	1045	56.2	374	4 US-09-442-349A-56	Sequence 56, Appli
44	1044	56.1	374	3 US-08-878-801-2	Sequence 2, Appli
45	1044	56.1	374	4 US-09-442-349A-39	Sequence 39, Appli

ALIGNMENTS

RESULT 1									
US-08-878-801-4									
Sequence 4, Application US/08878801									
Patent No. 6004808									
GENERAL INFORMATION:									
APPLICANT: Negulescu, Paul									
APPLICANT: Oftermanns, Stefan									
APPLICANT: Simon, Melvin									
APPLICANT: Zuker, Charles									
TITLE OF INVENTION: PROMISCUOUS G-PROTEINS COMPOSITIONS AND THEIR USE									
FILE REFERENCE: 08366/002001									
CURRENT APPLICATION NUMBER: US/08/878,801									
CURRENT FILING DATE: 1997-06-19									
EARLIER FILING DATE: 1996-06-21									
NUMBER OF SEQ ID NOS: 4									
SOFTWARE: FASTSEQ for Windows Version 3.0									
SEQ ID NO 4									
LENGTH: 374									
TYPE: PRT									
ORGANISM: Homo sapiens									
US-08-878-801-4									
Query Match									
Best Local Similarity 57.3%; Score 1066; DB 3; Length 374;									
Matches 207; Conservative 55; Mismatches 86; Indels 12; Gaps 2;									
QY	10	CLSEAKREARRINDEIRHVRDRKRDAREKLILLGTGSGSKSTFIKMRILHSGSYSD	69						
DB	13	CLTEERKNAIIDDEINILLGKOEHEELKLLLGSGSGSKSTFIKMRILHSGSYSD	72						
QY	70	EDKRGFTLVYONIFITAMQAMIRAMDTIKIDPKYEHKKAHOLVRENDVEKVSAPENDYV	129						
DB	73	EDBRAFRLLIYONIFITAMQAMIRAMDTIKIDPKYEHKKAHOLVRENDVEKVSAPENDYV	132						
QY	130	DAISLWMDPDIQOEYDRREYOLSDSKYKYLINDLRADAPAYLPTOODVLRVYPTTGI	189						
DB	133	VAMQYLDADGIRCYERREHFLHDSAVVYLISHLERSISDSYIPTADVLRSRMPTTGI	192						
QY	190	IEYFPDLOSIVFRMWDVGQSESRKRWTHCEENTYSIMFVLALESYDOLVESDNENME	249						
DB	193	MEYCFSVKTKLRIVDVGQSESRKRWTHCEENTYSIMFVLALESYDOLVESDNENME	252						
QY	250	ESKALFTIITVYPMFQNSVILFLNKKDLLEKIMYSHLVYFPYDGPQDQAAREFI	309						
DB	253	ESLAFSTILFLPMFKSTSVILFLNKKDLLEKIMYSHLVYFPYDGPQDQAAREFI	312						
QY	310	LKMEVDL-----NDSD-----KTIISHTCATDPMNIFVFAAVYDITLQLNKEYN	357						
DB	313	LDMYARVYASCAEPDGGGRKSRARRRFFAHTCAVSVRSVFKVDVSLARYLDEIN	372						

QY 358 LV 359  
1:  
Db 373 LL 374

## RESULT 2

US-09-442-349A-29  
; Sequence 29, Application US/09442349A  
; Patent No. 6462178  
; GENERAL INFORMATION:  
; APPLICANT: Mong, Yung H  
; TITLE OF INVENTION: G Protein  
; FILE REFERENCE: M99/0101/US  
; CURRENT APPLICATION NUMBER: US/09/442,349A  
; CURRENT FILING DATE: 1999-11-17  
; NUMBER OF SEQ ID NOS: 116  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 29  
; LENGTH: 374  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: G protein.  
US-09-442-349A-29

Query Match 57.2%; Score 1063; DB 4; Length 374;  
Best Local Similarity 57.1%; Pred. No. 1.6e-99;  
Matches 206; Conservative 57; Mismatches 86; Indels 12; Gaps 2;

QY 10 CLSEEAKEARRINDEIRRHVRDQKRDARRELKLLLTGSGSGSTFIKQRIIHGSGYSD 69  
13 CLTEDEKAARVQDEINRILLEOKKODRGELKLLLGPGSGSTFIKQRIIHGAGYSE 72  
Db 70 EDRKGFTKLYQNIFTQAMIRAMDCLKIPYKEHNKAHAOLVREVDKVSASFENPVY 129  
73 EERKGFRLVYQNIIFVSMRAMIEMERLQIPFSRPSKHHASLVMSODPKVTTFFEKRYA 132  
QY 130 DAISLWMDPGIOECYDRRREYOLSDSTKYLYLNDLRVADPAYLPQODVLRVAVPTGT 189  
133 AAMQWLMDRAGIRACYERREPHLDASVYVYLSHLERITEGVYPTQADVLRSMPTGT 192  
Db 190 IEYFDQSVIFRNVDVGGRSERKKIHCFENVTSIMFLVALSEYQVLESNDENRME 249  
193 NEYCFYQKTNLRIVDVGGRSERKKIHCFENVIALITYLASLSEYDQCLEENNQENRME 252  
QY 250 ESKALFRTITVPMFQNSVILFLNKKDLLEEKIMYSHLVDPPEYDGPORDAQAAREFI 309  
253 ESLALFTLTLELPMFKSTVILFLNKTDLLEKIPSHLATYPPSPGPRKQDAEAARKFI 312  
Db 310 LKMFVLDI-----NPSD-----KIYSHFTCATDTENIRVPAVAKDTILOLNKEYN 357  
313 LDMYTRMYTGCVDGPEGSSNNRKEKTEIYSHFTCATDTKNIQFVAVTDTVIIONNLKYIG 372  
QY 358 L 358  
Db 373 L 373

## RESULT 3

US-09-442-349A-31  
; Sequence 31, Application US/09442349A  
; Patent No. 6462178  
; GENERAL INFORMATION:  
; APPLICANT: Mong, Yung H  
; TITLE OF INVENTION: G Protein  
; FILE REFERENCE: M99/0101/US  
; CURRENT APPLICATION NUMBER: US/09/442,349A  
; CURRENT FILING DATE: 1999-11-17  
; NUMBER OF SEQ ID NOS: 116  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 31

LENGTH: 374  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: G protein  
US-09-442-349A-31

Query Match 57.2%; Score 1063; DB 4; Length 374;  
Best Local Similarity 57.1%; Pred. No. 1.6e-99;  
Matches 206; Conservative 57; Mismatches 86; Indels 12; Gaps 2;

QY 10 CLSEEAKEARRINDEIRRHVRDQKRDARRELKLLLTGSGSGSTFIKQRIIHGSGYSD 69  
13 CLTEDEKAARVQDEINRILLEOKKODRGELKLLLGPGSGSTFIKQRIIHGAGYSE 72  
Db 70 EDRKGFTKLYQNIFTQAMIRAMDCLKIPYKEHNKAHAOLVREVDKVSASFENPVY 129  
73 EERKGFRLVYQNIIFVSMRAMIEMERLQIPFSRPSKHHASLVMSODPKVTTFFEKRYA 132  
QY 130 DAISLWMDPGIOECYDRRREYOLSDSTKYLYLNDLRVADPAYLPQODVLRVAVPTGT 189  
133 AAMQWLMDRAGIRACYERREPHLDASVYVYLSHLERITEGVYPTQADVLRSMPTGT 192  
Db 190 IEYFDQSVIFRNVDVGGRSERKKIHCFENVTSIMFLVALSEYQVLESNDENRME 249  
193 NEYCFYQKTNLRIVDVGGRSERKKIHCFENVIALITYLASLSEYDQCLEENNQENRME 252  
QY 250 ESKALFRTITVPMFQNSVILFLNKKDLLEEKIMYSHLVDPPEYDGPORDAQAAREFI 309  
253 ESLALFTLTLELPMFKSTVILFLNKTDLLEKIPSHLATYPPSPGPRKQDAEAARKFI 312  
Db 310 LKMFVLDI-----NPSD-----KIYSHFTCATDTENIRVPAVAKDTILOLNKEYN 357  
313 LDMYTRMYTGCVDGPEGSSNNRKEKTEIYSHFTCATDTKNIQFVAVTDTVIIONNLKYIG 372  
QY 358 L 358  
Db 373 L 373

## RESULT 4

US-09-442-349A-21  
; Sequence 21, Application US/09442349A  
; Patent No. 6462178  
; GENERAL INFORMATION:  
; APPLICANT: Mong, Yung H  
; TITLE OF INVENTION: G Protein  
; FILE REFERENCE: M99/0101/US  
; CURRENT APPLICATION NUMBER: US/09/442,349A  
; CURRENT FILING DATE: 1999-11-17  
; NUMBER OF SEQ ID NOS: 116  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 21  
; LENGTH: 374  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: G protein  
US-09-442-349A-21

Query Match 57.1%; Score 1062; DB 4; Length 374;  
Best Local Similarity 57.1%; Pred. No. 2e-99;  
Matches 206; Conservative 57; Mismatches 86; Indels 12; Gaps 2;

QY 10 CLSEEAKEARRINDEIRRHVRDQKRDARRELKLLLTGSGSGSTFIKQRIIHGSGYSD 69  
13 CLTEDEKAARVQDEINRILLEOKKODRGELKLLLGPGSGSGSTFIKQRIIHGAGYSE 72  
Db 70 EDRKGFTKLYQNIFTQAMIRAMDCLKIPYKEHNKAHAOLVREVDKVSASFENPVY 129  
73 EERKGFRLVYQNIIFVSMRAMIEMERLQIPFSRPSKHHASLVMSODPKVTTFFEKRYA 132



```

DB      373 L 373

RESULT 6
US-09-442-349A-25
; Sequence 25, Application US/09442349A
; Patent No. 6462178
; GENERAL INFORMATION:
; APPLICANT: Mong, Yung H
; TITLE OF INVENTION: G Protein
; FILE REFERENCE: M99/0101/US
; CURRENT APPLICATION NUMBER: US/09/442,349A
; CURRENT FILING DATE: 1999-11-17
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 25
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: G protein
US-09-442-349A-25

Query Match          57.1%; Score 1062; DB 4; Length 374;
Best Local Similarity 57.1%; Pred. No. 2e-99;
Matches 206; Conservative 56; Mismatches 97; Indels 12; Gaps 2;

QY      10 CLSEFAKRRARINDEIERHVRDRKDKARRELKLLLTGSGSKSFNIKOMRIHSGSYSD 69
      11 |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
DB      13 CLTEDEKAAARVDQETIRNILEOKKQDNGELKLLGLGSGSKSTFIOMRIHAGATSE 72
      14 |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|

QY      70 EDRKGFRLVYQNIETAMQAMIRAMDTLKIPIKYEHKNAHAQLREVDVEKVSAPENPYV 129
      71 |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
DB      73 EERKGFRLVYQNIETAMQAMIRAMDTLKIPIKYEHKNAHAQLREVDVEKVSAPENPYV 132
      74 |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|

QY      130 DAISLMDPQIGOEYDPRREYQLSDSTKYIINDLRADRAYLLETQDYLRVRYPTTGI 189
      131 |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
DB      133 AAMQWLMDAGIRACYEYRRERREHLLDSAVVYLSHLERTEEYVPTADVDLSRMPPTGI 192
      134 |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|

QY      190 IEYPPDLDSVIFRWMDVGGQSRERKWHCFENYVSIHFLVALESDVLESDNENME 249
      191 |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
DB      193 NEYCSYQKTNRIYDVGGQSRERKWHCFENYVSIHFLVALESDVLESDNENME 252
      194 |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|

QY      250 ESKALFRITIIYPMFQNSVILFLKKDLLEEKIMYSHLVDFPEPYDGPQRDAQAREFI 309
      251 |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
DB      253 ESLAFGLTIIELPMFKSTSVILFLKKDTLLEEKIPTSHLATYFPFSGPKODAEAKRFI 312
      254 |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|

QY      310 LKMFVDL-----NDSD-----KIIYSHFCATDENIRVFVAAYVDTIQLCLKEN 357
      311 |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
DB      313 LDMYRMITTCVDPGEGSGNRKREKYEIYSHFTCAIDTISNIOVFQAVIDVLIIONNLK 372
      314 |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|

QY      358 L 358
DB      373 L 373

RESULT 7
US-09-442-349A-30
; Sequence 30, Application US/09442349A
; Patent No. 6462178
; GENERAL INFORMATION:
; APPLICANT: Mong, Yung H
; TITLE OF INVENTION: G Protein
; FILE REFERENCE: M99/0101/US
; CURRENT APPLICATION NUMBER: US/09/442,349A
; CURRENT FILING DATE: 1999-11-17
; NUMBER OF SEQ ID NOS: 116
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 30
; LENGTH: 374
; TYPE: PRT
; ORGANISM: Artificial Sequence

```

FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: G protein  
OTHER INFORMATION: chimera  
US-09-442-349A-30

Query Match 57.1%; Score 1062; DB 4; Length 374;  
Best Local Similarity 57.1%; Pred. No. 2e-99;  
Matches 206; Conservative 57; Mismatches 86; Indels 12; Gaps 2;

QY 10 CLSEAKARRINDEIERHVRDQRARRELKLLLGTSKSTFIKQRIIHGSGYSD 69  
DB 13 CLTEDEKAAARVQOEINRILLEOKKODRGELKLLLGPGSGSTFIKQRIIHGAGYSE 72  
QY 70 EDRKGFRTKLYQNIETAMQAMIRAMDPLKIPYKEHNKAHAOLVREVDVKSASFENPYV 129  
DB 73 EERKGFRTPLVYQNIETAMQAMIRAMDPLKIPYKEHNKAHAOLVREVDVKSASFENPYV 132  
QY 130 DAISLWNPDIQECYDRRREYQLSSTKYLLNDLRVADPAYLPTQODVLRVAVPTTGI 189  
DB 133 AAMQMLMRDAGIRACYERRRREFHLLDSAVYLLSHLERTEGYVPTADVLRSMPTTGI 192  
QY 190 IEYFDLOSIVFRWVGVGSGSRERKWIHCFENVTSMFLVALSEYDQVLVESDNENRME 249  
DB 193 NEYCFSVQKTNLRIVDVGSGSRERKWIHCFENVTSMFLVALSEYDQVLVESDNENRME 252  
QY 250 ESKALFRTIITYPMFQNSVILFLNKKDLLEEKIMYSHLVDPPEYDGPORDQAAREFI 309  
DB 253 ESIALFRTILELPMFKSTVILFLNKTDLLEEKIPTSHLATYFPFSGPKODAAAKRFI 312  
QY 310 LKMFVDL-----NPSD-----KIYSHFTCATDTENIRFVAVKDTIQLNLKEYN 357  
DB 313 LDMYTRMYTGCVDGPEGSNNRKEIKETIYSHFTCATDTNNIQFVAVTDTIQLNNLKYIG 372  
QY 358 L 358  
DB 373 L 373

RESULT 8  
US-09-442-349A-32  
Sequence 32, Application US/09442349A  
Patent No. 6462178  
GENERAL INFORMATION:  
APPLICANT: Wong, Yung H  
TITLE OF INVENTION: G Protein  
FILE REFERENCE: M99/0101/US  
CURRENT APPLICATION NUMBER: US/09/442.349A  
CURRENT FILING DATE: 1999-11-17  
NUMBER OF SEQ ID NOS: 116  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 32  
LENGTH: 374  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: G protein  
OTHER INFORMATION: chimera  
US-09-442-349A-32

Query Match 57.1%; Score 1062; DB 4; Length 374;  
Best Local Similarity 57.1%; Pred. No. 2e-99;  
Matches 206; Conservative 56; Mismatches 87; Indels 12; Gaps 2;

QY 10 CLSEAKARRINDEIERHVRDQRARRELKLLLGTSKSTFIKQRIIHGSGYSD 69  
DB 13 CLTEDEKAAARVQOEINRILLEOKKODRGELKLLLGPGSGSTFIKQRIIHGAGYSE 72  
QY 70 EDRKGFRTKLYQNIETAMQAMIRAMDPLKIPYKEHNKAHAOLVREVDVKSASFENPYV 129  
DB 73 EERKGFRTPLVYQNIETAMQAMIRAMDPLKIPYKEHNKAHAOLVREVDVKSASFENPYV 132  
QY 130 DAISLWNPDIQECYDRRREYQLSSTKYLLNDLRVADPAYLPTQODVLRVAVPTTGI 189

DB 133 AAMQMLMRDAGIRACYERRRREFHLLDSAVYLLSHLERTEGYVPTADVLRSMPTTGI 192  
QY 190 IEYFDLOSIVFRWVGVGSGSRERKWIHCFENVTSMFLVALSEYDQVLVESDNENRME 249  
DB 193 NEYCFSVQKTNLRIVDVGSGSRERKWIHCFENVTSMFLVALSEYDQVLVESDNENRME 252  
QY 250 ESKALFRTIITYPMFQNSVILFLNKKDLLEEKIMYSHLVDPPEYDGPORDQAAREFI 309  
DB 253 ESIALFRTILELPMFKSTVILFLNKTDLLEEKIPTSHLATYFPFSGPKODAAAKRFI 312  
QY 310 LKMFVDL-----NPSD-----KIYSHFTCATDTENIRFVAVKDTIQLNLKEYN 357  
DB 313 LDMYTRMYTGCVDGPEGSNNRKEIKETIYSHFTCATDTNNIQFVAVTDTIQLNNLKYIG 372  
QY 358 L 358  
DB 373 L 373

RESULT 9  
US-09-442-349A-34  
Sequence 34, Application US/09442349A  
Patent No. 6462178  
GENERAL INFORMATION:  
APPLICANT: Wong, Yung H  
TITLE OF INVENTION: G Protein  
FILE REFERENCE: M99/0101/US  
CURRENT APPLICATION NUMBER: US/09/442.349A  
CURRENT FILING DATE: 1999-11-17  
NUMBER OF SEQ ID NOS: 116  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 34  
LENGTH: 374  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: G protein  
OTHER INFORMATION: chimera  
US-09-442-349A-34

Query Match 57.1%; Score 1062; DB 4; Length 374;  
Best Local Similarity 57.1%; Pred. No. 2e-99;  
Matches 206; Conservative 56; Mismatches 87; Indels 12; Gaps 2;

QY 10 CLSEAKARRINDEIERHVRDQRARRELKLLLGTSKSTFIKQRIIHGSGYSD 69  
DB 13 CLTEDEKAAARVQOEINRILLEOKKODRGELKLLLGPGSGSTFIKQRIIHGAGYSE 72  
QY 70 EDRKGFRTKLYQNIETAMQAMIRAMDPLKIPYKEHNKAHAOLVREVDVKSASFENPYV 129  
DB 73 EERKGFRTPLVYQNIETAMQAMIRAMDPLKIPYKEHNKAHAOLVREVDVKSASFENPYV 132  
QY 130 DAISLWNPDIQECYDRRREYQLSSTKYLLNDLRVADPAYLPTQODVLRVAVPTTGI 189  
DB 133 AAMQMLMRDAGIRACYERRRREFHLLDSAVYLLSHLERTEGYVPTADVLRSMPTTGI 192  
QY 190 IEYFDLOSIVFRWVGVGSGSRERKWIHCFENVTSMFLVALSEYDQVLVESDNENRME 249  
DB 193 NEYCFSVQKTNLRIVDVGSGSRERKWIHCFENVTSMFLVALSEYDQVLVESDNENRME 252  
QY 250 ESKALFRTIITYPMFQNSVILFLNKKDLLEEKIMYSHLVDPPEYDGPORDQAAREFI 309  
DB 253 ESIALFRTILELPMFKSTVILFLNKTDLLEEKIPTSHLATYFPFSGPKODAAAKRFI 312  
QY 310 LKMFVDL-----NPSD-----KIYSHFTCATDTENIRFVAVKDTIQLNLKEYN 357  
DB 313 LDMYTRMYTGCVDGPEGSNNRKEIKETIYSHFTCATDTNNIQFVAVTDTIQLNNLKYIG 372  
QY 358 L 358  
DB 373 L 373



```
Db      193  NEXCFSVQKTNLRIVDVGQSKSEKKWIHCFENYIALIYLAISEYDQCLEENNOENRMK 252
      250  ESKALFRTITTYPFQWONSSVILFLNKDLEEKIMYSHLVYPEYDGPORDAQAAREFI 309
      Db      253  ESKALFRTITTYPFQWONSSVILFLNKDLEEKIMYSHLVYPEYDGPORDAQAAREFI 312
      Qy      310  LKMFVVDL-----NPSD-----KIYSHFTCATDPTENIRFVFAVKDTLLQNLKEYN 357
      Db      313  LDMYTRMYTGCVDGPEGSNNRKKETKIYSHFTCATDPTENIRFVFAVKDTLLQNLKEYN 372
      Qy      358  L 358
      Db      373  L 373
```

## RESULT 13

US-09-442-349A-33

```
Sequence 33: Application US/09442349A
Patent No. 6462178
GENERAL INFORMATION:
APPLICANT: Mong, Yung H
TITLE OF INVENTION: G Protein
FILE REFERENCE: M99/0101/US
CURRENT APPLICATION NUMBER: US/09/442,349A
NUMBER OF SEQ ID NOS: 116
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 33
LENGTH: 374
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: G protein
US-09-442-349A-33
```

Query Match 57.0%; Score 1061; DB 4; Length 374;

Best Local Similarity 57.1%; Pred. No. 2,5e-99;

Matches 206; Conservative 56; Mismatches 87; Indels 12; Gaps 2;

```
Qy      10  CLSEAKERRINDEIRHYARDKRDARRELKLLLGTSKGSTFTKOMRIHSGSYD 69
      Db      13  CLTDEKAARVDEIRIRILLEOKKODRGELKLLLPGESGKSTFTKOMRIHAGAYSE 72
      Qy      70  EDKRGFTKLYQNIPTAMQAMIRAMDLPKPYKXENKHAQAOLYREVDEKVSAPENPYV 129
      Db      73  EERKGFRLPYQNIPTFVSMRAMIEMERLQIPFSRPSKHNASLVMSODPYKVTTFEKKRYA 132
      Qy      130  DAISLWNDPGIOECYRRREYQSLDSTKYLLNDLVADPAVLPFOODVLRAVPTGTI 189
      Db      133  AAMQMLRDAIGIRACVRRREFHLLDSAVYYLSHLERITEGYVPTAQDVLRSMPPTGI 192
      Qy      190  IEYFPDQSVIFRWVVDVGQSGRSEKKWIHCFENYTSIMPLVAISEYDQVLESNENRME 249
      Db      193  NEXCFSVQKTNLRIVDVGQSKSEKKWIHCFENYIALIYLAISEYDQCLEENNOENRMK 252
      Qy      250  ESKALFRTITTYPFQWONSSVILFLNKDLEEKIMYSHLVYPEYDGPORDAQAAREFI 309
      Db      253  ESKALFRTITTYPFQWONSSVILFLNKDLEEKIMYSHLVYPEYDGPORDAQAAREFI 312
      Qy      310  LKMFVVDL-----NPSD-----KIYSHFTCATDPTENIRFVFAVKDTLLQNLKEYN 357
      Db      313  LDMYTRMYTGCVDGPEGSNNRKKETKIYSHFTCATDPTENIRFVFAVKDTLLQNLKEYN 372
      Qy      358  L 358
      Db      373  L 373
```

RESULT 14  
US-09-442-349A-26  
Sequence 26: Application US/09442349A

```
Patent No. 6462178
GENERAL INFORMATION:
APPLICANT: Mong, Yung H
TITLE OF INVENTION: G Protein
FILE REFERENCE: M99/0101/US
CURRENT APPLICATION NUMBER: US/09/442,349A
NUMBER OF SEQ ID NOS: 116
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 26
LENGTH: 374
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: G protein
US-09-442-349A-26
```

Query Match 56.9%; Score 1058; DB 4; Length 374;

Best Local Similarity 57.1%; Pred. No. 5e-99;

Matches 206; Conservative 56; Mismatches 87; Indels 12; Gaps 2;

```
Qy      10  CLSEAKERRINDEIRHYARDKRDARRELKLLLGTSKGSTFTKOMRIHSGSYD 69
      Db      13  CLTDEKAARVDEIRIRILLEOKKODRGELKLLLPGESGKSTFTKOMRIHAGAYSE 72
      Qy      70  EDKRGFTKLYQNIPTAMQAMIRAMDLPKPYKXENKHAQAOLYREVDEKVSAPENPYV 129
      Db      73  EERKGFRLPYQNIPTFVSMRAMIEMERLQIPFSRPSKHNASLVMSODPYKVTTFEKKRYA 132
      Qy      130  DAISLWNDPGIOECYRRREYQSLDSTKYLLNDLVADPAVLPFOODVLRAVPTGTI 189
      Db      133  AAMQMLRDAIGIRACVRRREFHLLDSAVYYLSHLERITEGYVPTAQDVLRSMPPTGI 192
      Qy      190  IEYFPDQSVIFRWVVDVGQSGRSEKKWIHCFENYTSIMPLVAISEYDQVLESNENRME 249
      Db      193  NEXCFSVQKTNLRIVDVGQSKSEKKWIHCFENYIALIYLAISEYDQCLEENNOENRMK 252
      Qy      250  ESKALFRTITTYPFQWONSSVILFLNKDLEEKIMYSHLVYPEYDGPORDAQAAREFI 309
      Db      253  ESKALFRTITTYPFQWONSSVILFLNKDLEEKIMYSHLVYPEYDGPORDAQAAREFI 312
      Qy      310  LKMFVVDL-----NPSD-----KIYSHFTCATDPTENIRFVFAVKDTLLQNLKEYN 357
      Db      313  LDMYTRMYTGCVDGPEGSNNRKKETKIYSHFTCATDPTENIRFVFAVKDTLLQNLKEYN 372
      Qy      358  L 358
      Db      373  L 373
```

## RESULT 15

US-09-442-349A-28

```
Sequence 28: Application US/09442349A
Patent No. 6462178
GENERAL INFORMATION:
APPLICANT: Mong, Yung H
TITLE OF INVENTION: G Protein
FILE REFERENCE: M99/0101/US
CURRENT APPLICATION NUMBER: US/09/442,349A
NUMBER OF SEQ ID NOS: 116
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 28
LENGTH: 374
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: G protein
US-09-442-349A-28
```

Query Match 56.9%; Score 1058; DB 4; Length 374;





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2003, 16:16:21 ; Search time 24 Seconds  
(without alignments)  
1640.134 Million cell updates/sec

Title: US-09-899-295-2  
Perfect score: 1860  
Sequence: 1 MTEISIMACCLSEAKRARR.....VFAAVDTILQLNKEYNLV 359

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 424699 seqs, 109646833 residues

Total number of hits satisfying chosen parameters: 424699

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PTCT\_NEW\_PUB pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/PTCTUS\_PUBCOMB pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB pep:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB pep:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1860	100.0	359	10	US-09-899-295-2
2	1846	99.2	359	10	US-09-984-292-1
3	1846	99.2	359	10	US-09-989-497-1
4	1842	99.0	359	10	US-09-984-292-37
5	1842	99.0	359	10	US-09-989-497-37
6	1837	98.8	359	10	US-09-984-292-15
7	1837	98.8	359	10	US-09-989-497-15
8	1822	98.0	359	10	US-09-984-292-9
9	1822	98.0	359	10	US-09-989-497-9
10	1819	97.8	353	10	US-09-984-292-2
11	1819	97.8	353	10	US-09-989-497-2
12	1818	97.7	359	10	US-09-984-292-21
13	1818	97.7	359	10	US-09-989-497-21
14	1817	97.7	353	9	US-09-952-680A-27
15	1817	97.7	353	10	US-09-984-292-16
16	1817	97.7	353	10	US-09-989-497-16
17	1815	97.6	359	10	US-09-984-292-11
18	1815	97.6	359	10	US-09-989-497-11
19	1814	97.5	353	10	US-09-899-295-6

20	1814	97.5	353	10	US-09-899-295-8	Sequence 8, App1
21	1813	97.5	359	10	US-09-984-292-26	Sequence 26, App1
22	1813	97.5	359	10	US-09-989-497-26	Sequence 26, App1
23	1811	97.4	359	10	US-09-984-292-3	Sequence 3, App1
24	1811	97.4	359	10	US-09-989-497-3	Sequence 3, App1
25	1810	97.3	359	10	US-09-899-295-4	Sequence 20, App1
26	1807	97.2	359	10	US-09-984-292-20	Sequence 20, App1
27	1807	97.2	359	10	US-09-989-497-20	Sequence 20, App1
28	1806	97.1	359	10	US-09-984-292-25	Sequence 25, App1
29	1806	97.1	359	10	US-09-989-497-25	Sequence 25, App1
30	1804	97.0	359	10	US-09-984-292-10	Sequence 10, App1
31	1804	97.0	359	10	US-09-989-497-10	Sequence 10, App1
32	1803	96.9	359	10	US-09-984-292-22	Sequence 22, App1
33	1803	96.9	359	10	US-09-989-497-22	Sequence 22, App1
34	1800	96.8	359	10	US-09-984-292-12	Sequence 12, App1
35	1800	96.8	359	10	US-09-989-497-12	Sequence 12, App1
36	1795	96.5	353	10	US-09-984-292-8	Sequence 8, App1
37	1795	96.5	353	10	US-09-989-497-8	Sequence 8, App1
38	1793	96.4	359	10	US-09-984-292-13	Sequence 13, App1
39	1793	96.4	359	10	US-09-989-497-13	Sequence 13, App1
40	1791	96.3	353	10	US-09-984-292-17	Sequence 17, App1
41	1791	96.3	353	10	US-09-989-497-17	Sequence 17, App1
42	1789	96.2	359	10	US-09-984-292-23	Sequence 23, App1
43	1789	96.2	359	10	US-09-989-497-23	Sequence 23, App1
44	1784	95.9	353	10	US-09-984-292-4	Sequence 4, App1
45	1784	95.9	353	10	US-09-984-292-5	Sequence 5, App1

ALIGNMENTS

RESULT 1  
US-09-899-295-2  
Sequence 2, Application US/09899295  
Patent No.: US20020127601A1  
GENERAL INFORMATION:  
APPLICANT: Aventis Pharma Deutschland GmbH  
TITLE OF INVENTION: Process for identifying modulators of G protein coupled  
FILE REFERENCE: AVE D-2000/A033 englisch  
CURRENT APPLICATION NUMBER: US/09/899,295  
CURRENT FILING DATE: 2001-07-06  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 359  
TYPE: PRT  
ORGANISM: Mus musculus  
US-09-899-295-2

Query Match	100.0%;	Score 1860;	DB 10;	Length 359;
Best Local Similarity	100.0%;	Pred. No. 1.8e-149;		
Matches 359;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MTEISIMACCLSEAKRARRINDEIERHVRDRKRDARREKLILGTGSGSKSTFIKMR	60	
DB	1	MTEISIMACCLSEAKRARRINDEIERHVRDRKRDARREKLILGTGSGSKSTFIKMR	60	
QY	61	IINSGYSDDKRGFTLVYQNIPTAMQAMIRADTIKIYKKEHNAHQVREVDYER	120	
DB	61	IINSGYSDDKRGFTLVYQNIPTAMQAMIRADTIKIYKKEHNAHQVREVDYER	120	
QY	121	VSAPENYVYAIKSLMDPGIOECYDRRREYQSDSTKYYLLNDPVRADPAVPTQDVL	180	
DB	121	VSAPENYVYAIKSLMDPGIOECYDRRREYQSDSTKYYLLNDPVRADPAVPTQDVL	180	
QY	181	RVRVPTGIIIEYPPDLOSIVFRMYDVGGSERRKWIHCEENWTSIMFLVALSEYDOVL	240	
DB	181	RVRVPTGIIIEYPPDLOSIVFRMYDVGGSERRKWIHCEENWTSIMFLVALSEYDOVL	240	
QY	241	ESDNEHMEKSKALFRITIIYVFWFQNSVILFLNKKLLEKIMYSHLVYFPYDGPOR	300	
DB	241	ESDNEHMEKSKALFRITIIYVFWFQNSVILFLNKKLLEKIMYSHLVYFPYDGPOR	300	

OY 301 DAQAAREFLIKMFVDLNPDSKIIYSHFTCATDTENIRFVAVKDTIILQNLKEYNLV 359  
Db 301 DAQAAREFLIKMFVDLNPDSKIIYSHFTCATDTENIRFVAVKDTIILQNLKEYNLV 359

## RESULT 2

US-09-984-292-1  
; Sequence 1, Application US/09984292  
; Patent No. US20020128433A1  
; GENERAL INFORMATION:  
; APPLICANT: YAO, YONG  
; APPLICANT: XU, HONG

; TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE  
; TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF  
; FILE REFERENCE: 078003-0280649  
; CURRENT APPLICATION NUMBER: US/09/984,292  
; PRIOR APPLICATION NUMBER: 60/243,770  
; PRIOR FILING DATE: 2000-10-30  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 359  
; TYPE: PRT  
; ORGANISM: Mus sp.

US-09-984-292-1

Query Match 99.2%; Score 1846; DB 10; Length 359;  
Best Local Similarity 99.2%; Pred. No. 2,7e-148;

Matches 356; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 MLESIMACCLSEAEKARRINDEIERHVRDRKRDARRELKLLLTGSGSKSTFIKQMR 60  
Db 1 MLESIMACCLSEAEKARRINDEIERHVRDRKRDARRELKLLLTGSGSKSTFIKQMR 60  
OY 61 IIHSGSYSDCKRGFTKLYQNIPTAMQAMIRAMDITKIPYKKEHNAHQVREVYDVK 120  
Db 61 IIHSGSYSDCKRGFTKLYQNIPTAMQAMIRAMDITKIPYKKEHNAHQVREVYDVK 120  
OY 121 VSAFENPYDAIKSLMNDPQIOECYDRRREYQSDSTKYYLNDLDRVADPAYLPTQODVL 180  
Db 121 VSAFENPYDAIKSLMNDPQIOECYDRRREYQSDSTKYYLNDLDRVADPAYLPTQODVL 180  
OY 181 RVAVPTGIIIEYFPDQSVIFRWDVGQSRERKWHCEENVTSIMFLVALSEYDQVLY 240  
Db 181 RVAVPTGIIIEYFPDQSVIFRWDVGQSRERKWHCEENVTSIMFLVALSEYDQVLY 240  
OY 241 ESNENNMESKALFRTIITYPWFONSSVLLFLNKKDLLEKIMYSHLVDPPEYDGPOR 300  
Db 241 ESNENNMESKALFRTIITYPWFONSSVLLFLNKKDLLEKIMYSHLVDPPEYDGPOR 300  
OY 301 DAQAAREFLIKMFVDLNPDSKIIYSHFTCATDTENIRFVAVKDTIILQNLKEYNLV 359  
Db 301 DAQAAREFLIKMFVDLNPDSKIIYSHFTCATDTENIRFVAVKDTIILQNLKEYNLV 359

## RESULT 3

US-09-989-497-1  
; Sequence 1, Application US/09989497  
; Patent No. US20020143151A1  
; GENERAL INFORMATION:  
; APPLICANT: YAO, YONG  
; APPLICANT: XU, HONG

; TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE  
; TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF  
; FILE REFERENCE: 078003-0280735  
; CURRENT APPLICATION NUMBER: US/09/989,497  
; CURRENT FILING DATE: 2001-11-21  
; PRIOR APPLICATION NUMBER: 09/984,292  
; PRIOR FILING DATE: 2001-10-29

; PRIOR APPLICATION NUMBER: 60/243,770  
; PRIOR FILING DATE: 2000-10-30  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 359  
; TYPE: PRT  
; ORGANISM: Mus sp.

US-09-989-497-1

Query Match 99.2%; Score 1846; DB 10; Length 359;  
Best Local Similarity 99.2%; Pred. No. 2,7e-148;  
Matches 356; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 MLESIMACCLSEAEKARRINDEIERHVRDRKRDARRELKLLLTGSGSKSTFIKQMR 60  
Db 1 MLESIMACCLSEAEKARRINDEIERHVRDRKRDARRELKLLLTGSGSKSTFIKQMR 60  
OY 61 IIHSGSYSDCKRGFTKLYQNIPTAMQAMIRAMDITKIPYKKEHNAHQVREVYDVK 120  
Db 61 IIHSGSYSDCKRGFTKLYQNIPTAMQAMIRAMDITKIPYKKEHNAHQVREVYDVK 120  
OY 121 VSAFENPYDAIKSLMNDPQIOECYDRRREYQSDSTKYYLNDLDRVADPAYLPTQODVL 180  
Db 121 VSAFENPYDAIKSLMNDPQIOECYDRRREYQSDSTKYYLNDLDRVADPAYLPTQODVL 180  
OY 181 RVAVPTGIIIEYFPDQSVIFRWDVGQSRERKWHCEENVTSIMFLVALSEYDQVLY 240  
Db 181 RVAVPTGIIIEYFPDQSVIFRWDVGQSRERKWHCEENVTSIMFLVALSEYDQVLY 240  
OY 241 ESNENNMESKALFRTIITYPWFONSSVLLFLNKKDLLEKIMYSHLVDPPEYDGPOR 300  
Db 241 ESNENNMESKALFRTIITYPWFONSSVLLFLNKKDLLEKIMYSHLVDPPEYDGPOR 300  
OY 301 DAQAAREFLIKMFVDLNPDSKIIYSHFTCATDTENIRFVAVKDTIILQNLKEYNLV 359  
Db 301 DAQAAREFLIKMFVDLNPDSKIIYSHFTCATDTENIRFVAVKDTIILQNLKEYNLV 359

## RESULT 4

US-09-984-292-37  
; Sequence 37, Application US/09984292  
; Patent No. US20020128433A1  
; GENERAL INFORMATION:  
; APPLICANT: YAO, YONG  
; APPLICANT: XU, HONG

; TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE  
; TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF  
; FILE REFERENCE: 078003-0280649  
; CURRENT APPLICATION NUMBER: US/09/984,292  
; CURRENT FILING DATE: 2001-10-29  
; PRIOR APPLICATION NUMBER: 60/243,770  
; PRIOR FILING DATE: 2000-10-30  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 37  
; LENGTH: 359  
; TYPE: PRT  
; ORGANISM: Homo sapiens

US-09-984-292-37

Query Match 99.0%; Score 1842; DB 10; Length 359;  
Best Local Similarity 99.2%; Pred. No. 5.9e-148;  
Matches 356; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 MLESIMACCLSEAEKARRINDEIERHVRDRKRDARRELKLLLTGSGSKSTFIKQMR 60  
Db 1 MLESIMACCLSEAEKARRINDEIERHVRDRKRDARRELKLLLTGSGSKSTFIKQMR 60  
OY 61 IIHSGSYSDCKRGFTKLYQNIPTAMQAMIRAMDITKIPYKKEHNAHQVREVYDVK 120  
Db 61 IIHSGSYSDCKRGFTKLYQNIPTAMQAMIRAMDITKIPYKKEHNAHQVREVYDVK 120





ORGANISM: Homo sapiens  
US-09-989-497-15  
Query Match 98.8%; Score 1837; DB 10; Length 359;  
Best Local Similarity 98.9%; Pred. No. 1.6e-147;  
Matches 355; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 1 MTLESIMACCLSEAKERRINDEIRHVRKRDARRELKLLLTGSGSKSTFIKQMR 60  
DB 1 MTLESIMACCLSEAKERRINDEIRHVRKRDARRELKLLLTGSGSKSTFIKQMR 60  
QY 61 ITHSGSYDEDKRGFTKLYQNIPTAMQAMIRAMDLPKIPYKEHNKAHAQVREVDER 120  
DB 61 ITHSGSYDEDKRGFTKLYQNIPTAMQAMIRAMDLPKIPYKEHNKAHAQVREVDER 120  
QY 121 VSAFENPYDAIKSLMNDPGIOECYDRRREYQSLDSTKYYLNDLDRVADPAYLPTQODVL 180  
DB 121 VSAFENPYDAIKSLMNDPGIOECYDRRREYQSLDSTKYYLNDLDRVADPAYLPTQODVL 180  
QY 181 RVRVPTGTIIEYFPDLOSIFRWDVGQSRERKWHCFENVTSMFLVALSEYDQVLY 240  
DB 181 RVRVPTGTIIEYFPDLOSIFRWDVGQSRERKWHCFENVTSMFLVALSEYDQVLY 240  
QY 241 ESDNENRMEESKALFTITITYPFONSSVILFLNKKDLLEKIMYSHLVDPPEYDGPQR 300  
DB 241 ESDNENRMEESKALFTITITYPFONSSVILFLNKKDLLEKIMYSHLVDPPEYDGPQR 300  
QY 301 DAOAAREFLIKMFVDLNPDSDKIITSHFTCATDTENIRFVAVKDTIQLNKEYNLV 359  
DB 301 DAOAAREFLIKMFVDLNPDSDKIITSHFTCATDTENIRFVAVKDTIQLNKEYNLV 359  
RESULT 8  
US-09-984-292-9  
Sequence 9, Application US/09984292  
Patent No. US20020128433A1  
GENERAL INFORMATION:  
APPLICANT: YAO, YONG  
APPLICANT: XU, HONG  
TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE  
TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF  
FILE REFERENCE: 078003-0280649  
CURRENT APPLICATION NUMBER: US/09/984,292  
PRIOR FILING DATE: 2001-10-29  
PRIOR APPLICATION NUMBER: 60/243,770  
NUMBER OF SEQ ID NOS: 42  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 9  
LENGTH: 359  
TYPE: PRT  
ORGANISM: Mus sp.  
US-09-984-292-9  
Query Match 98.0%; Score 1822; DB 10; Length 359;  
Best Local Similarity 98.3%; Pred. No. 2.9e-146;  
Matches 353; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
QY 1 MTLESIMACCLSEAKERRINDEIRHVRKRDARRELKLLLTGSGSKSTFIKQMR 60  
DB 1 MTLESIMACCLSEAKERRINDEIRHVRKRDARRELKLLLTGSGSKSTFIKQMR 60  
QY 61 ITHSGSYDEDKRGFTKLYQNIPTAMQAMIRAMDLPKIPYKEHNKAHAQVREVDER 120  
DB 61 ITHSGSYDEDKRGFTKLYQNIPTAMQAMIRAMDLPKIPYKEHNKAHAQVREVDER 120  
QY 121 VSAFENPYDAIKSLMNDPGIOECYDRRREYQSLDSTKYYLNDLDRVADPAYLPTQODVL 180  
DB 121 VSAFENPYDAIKSLMNDPGIOECYDRRREYQSLDSTKYYLNDLDRVADPAYLPTQODVL 180  
QY 181 RVRVPTGTIIEYFPDLOSIFRWDVGQSRERKWHCFENVTSMFLVALSEYDQVLY 240  
DB 181 RVRVPTGTIIEYFPDLOSIFRWDVGQSRERKWHCFENVTSMFLVALSEYDQVLY 240

DB 181 RVRVPTGTIIEYFPDLOSIFRWDVGQSRERKWHCFENVTSMFLVALSEYDQVLY 240  
QY 241 ESDNENRMEESKALFTITITYPFONSSVILFLNKKDLLEKIMYSHLVDPPEYDGPQR 300  
DB 241 ESDNENRMEESKALFTITITYPFONSSVILFLNKKDLLEKIMYSHLVDPPEYDGPQR 300  
QY 301 DAOAAREFLIKMFVDLNPDSDKIITSHFTCATDTENIRFVAVKDTIQLNKEYNLV 359  
DB 301 DAOAAREFLIKMFVDLNPDSDKIITSHFTCATDTENIRFVAVKDTIQLNKEYNLV 359  
RESULT 9  
US-09-989-497-9  
Sequence 9, Application US/09989497  
Patent No. US20020143151A1  
GENERAL INFORMATION:  
APPLICANT: YAO, YONG  
APPLICANT: XU, HONG  
TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE  
TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF  
FILE REFERENCE: 078003-0280735  
CURRENT APPLICATION NUMBER: US/09/989,497  
PRIOR FILING DATE: 2001-11-21  
PRIOR APPLICATION NUMBER: 09/984,292  
PRIOR FILING DATE: 2001-10-29  
PRIOR APPLICATION NUMBER: 60/243,770  
NUMBER OF SEQ ID NOS: 42  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 9  
LENGTH: 359  
TYPE: PRT  
ORGANISM: Mus sp.  
US-09-989-497-9  
Query Match 98.0%; Score 1822; DB 10; Length 359;  
Best Local Similarity 98.3%; Pred. No. 2.9e-146;  
Matches 353; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
QY 1 MTLESIMACCLSEAKERRINDEIRHVRKRDARRELKLLLTGSGSKSTFIKQMR 60  
DB 1 MTLESIMACCLSEAKERRINDEIRHVRKRDARRELKLLLTGSGSKSTFIKQMR 60  
QY 61 ITHSGSYDEDKRGFTKLYQNIPTAMQAMIRAMDLPKIPYKEHNKAHAQVREVDER 120  
DB 61 ITHSGSYDEDKRGFTKLYQNIPTAMQAMIRAMDLPKIPYKEHNKAHAQVREVDER 120  
QY 121 VSAFENPYDAIKSLMNDPGIOECYDRRREYQSLDSTKYYLNDLDRVADPAYLPTQODVL 180  
DB 121 VSAFENPYDAIKSLMNDPGIOECYDRRREYQSLDSTKYYLNDLDRVADPAYLPTQODVL 180  
QY 181 RVRVPTGTIIEYFPDLOSIFRWDVGQSRERKWHCFENVTSMFLVALSEYDQVLY 240  
DB 181 RVRVPTGTIIEYFPDLOSIFRWDVGQSRERKWHCFENVTSMFLVALSEYDQVLY 240  
QY 241 ESDNENRMEESKALFTITITYPFONSSVILFLNKKDLLEKIMYSHLVDPPEYDGPQR 300  
DB 241 ESDNENRMEESKALFTITITYPFONSSVILFLNKKDLLEKIMYSHLVDPPEYDGPQR 300  
QY 301 DAOAAREFLIKMFVDLNPDSDKIITSHFTCATDTENIRFVAVKDTIQLNKEYNLV 359  
DB 301 DAOAAREFLIKMFVDLNPDSDKIITSHFTCATDTENIRFVAVKDTIQLNKEYNLV 359  
RESULT 10  
US-09-984-292-2  
Sequence 2, Application US/09984292  
Patent No. US20020128433A1  
GENERAL INFORMATION:  
APPLICANT: YAO, YONG  
APPLICANT: XU, HONG  
TITLE OF INVENTION: G-ALPHA-Q PROTEIN VARIANTS AND THEIR USE IN THE

Query Match	97.8%	Score 1819	DB 10	Length 355
Best Local Similarity	99.28%	Pred. No. 5.1e-146		
Matches 350; Conservative	2;	Mismatches 1;	Indels 0;	Gaps 0

QY	1	MTLESIMACCLSEEAKEARINDEIETERNHVRPDKRDARREKLKLLLTGGSGKSTFKOMR	60
Db	1	MTLESIMACCLSEEAKEARINDEIETERNHVRPDKRDARREKLKLLLTGGSGKSTFKOMR	60
QY	61	IITHSGSYADEDKRGFTKLYVONIETAMQAMIRAMDTLKIPYKEHHKAHAQIVREYDVEK	1200
Db	61	IITHSGDYSEDDEKRGCTKLYVONITFAMQAMIRAMDTLKIPYKEHHKAHAQIVREYDVEK	1200
QY	121	VSASFENPYVADAKSLIMNDPGIOECYDRREYOLSDSTKYYLNDLDBVADPATYLPTOODYL	1800
QY	121	VSASFVPDYAAKSLIMNDPGIOECYDRREYOLSDSTKYYLNDLDBVADPATYLPTOODYL	1800
Db	121	VSASFVPDYAAKSLIMNDPGIOECYDRREYOLSDSTKYYLNDLDBVADPATYLPTOODYL	1800
QY	181	RVRVETPTGIIIEYPFDLOSVIFRMVDVGGGSRSEKKNIHCFEYNTSIMPVLVAISEYDOVLV	2400
Db	181	RVRVETPTGIIIEYPFDLOSVIFRMVDVGGGSRSEKKNIHCFEYNTSIMPVLVAISEYDOVLV	2400
QY	241	ESDNNRMBESKALPRTIITTPYMPFONSSVITLKNKDLLEBKIMYSHLVDPPEYGGPOR	3000
Db	241	ESDNNRMBESKALPRTIITTPYMPFONSSVITLKNKDLLEBKIMYSHLVDPPEYGGPOR	3000
QY	301	DAQAAREFLKMFVLDNPDSDKILYSHFCATDTEINIREVAAVKDTIILQNLNKEYNLV	359

Db 301 DAQAAREFLLKMFVDLNPDSDKIIYSHFTCATDTENIRFVAAVKDTIQLNLKEYNLV 359

## RESULT 13

US-09-989-497-21

; Sequence 21, Application US/09989497  
; Patent No. US20020143151A1  
; GENERAL INFORMATION:  
; APPLICANT: YAO, YONG  
; APPLICANT: XU, HONG  
; TITLE OF INVENTION: G-ALPHA-O PROTEIN VARIANTS AND THEIR USE IN THE  
; TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF  
; FILE REFERENCE: 078003-0280735  
; CURRENT APPLICATION NUMBER: US/09/989,497  
; PRIOR FILING DATE: 2001-11-21  
; PRIOR APPLICATION NUMBER: 09/984,292  
; PRIOR FILING DATE: 2001-10-29  
; PRIOR APPLICATION NUMBER: 60/243,770  
; PRIOR FILING DATE: 2000-10-30  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 21  
; LENGTH: 359  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-989-497-21

Query Match 97.7%; Score 1818; DB 10; Length 359;  
Best Local Similarity 98.3%; Pred. No. 6.3e-146;  
Matches 353; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MTELSIMACCLSEBAKARRINDEIRHVRDKRDARRELKLLLGSGSGSTFIKQMR 60  
DB 1 MTELSIMACCLSEBAKARRINDEIRHVRDKRDARRELKLLLGSGSGSTFIKQMR 60  
QY 61 ITHGSDYSDDEKRGFTKLVYQNIPTAMQAMIRAMDITKIPYKHNKAHQALREVYVER 120  
DB 61 ITHGSDYSDDEKRGFTKLVYQNIPTAMQAMIRAMDITKIPYKHNKAHQALREVYVER 120  
QY 121 VSAFENFYDAIKSLMNDPGIOECYDRRREYQLSDSTKYYLNDLDRVADPAVLPQODVL 180  
DB 121 VSAFENFYDAIKSLMNDPGIOECYDRRREYQLSDSTKYYLNDLDRVADPAVLPQODVL 180  
QY 181 RVAVPTGIIIEYFPDLOSIFRNVAVDGGORSEKRMHCEENVTSMFLVALSEYDQVLY 240  
DB 181 RVAVPTGIIIEYFPDLOSIFRNVAVDGGORSEKRMHCEENVTSMFLVALSEYDQVLY 240  
QY 241 ESDNENRMEESKALFTIITYPMFONSSVLLFLNKKDLLEKIMYSHLVDFPEYDGPQR 300  
DB 241 ESDNENRMEESKALFTIITYPMFONSSVLLFLNKKDLLEKIMYSHLVDFPEYDGPQR 300  
QY 241 ESDNENRMEESKALFTIITYPMFONSSVLLFLNKKDLLEKIMYSHLVDFPEYDGPQR 300  
DB 241 ESDNENRMEESKALFTIITYPMFONSSVLLFLNKKDLLEKIMYSHLVDFPEYDGPQR 300  
QY 301 DAQAAREFLLKMFVDLNPDSDKIIYSHFTCATDTENIRFVAAVKDTIQLNLKEYNLV 359  
DB 301 DAQAAREFLLKMFVDLNPDSDKIIYSHFTCATDTENIRFVAAVKDTIQLNLKEYNLV 359

## RESULT 14

US-09-952-680A-27

; Sequence 27, Application US/09952680A  
; Publication No. US20030087239A1  
; GENERAL INFORMATION:  
; APPLICANT: Stanton, Marty  
; APPLICANT: Epstein, David  
; APPLICANT: Hamaguchi, No. US20030087239A1uko  
; TITLE OF INVENTION: Target Activated Biosensor and Methods of Using Same  
; FILE REFERENCE: 23239-501  
; CURRENT APPLICATION NUMBER: US/09/952,680A  
; PRIOR FILING DATE: 2001-09-13  
; PRIOR APPLICATION NUMBER: 60/232,454  
; PRIOR FILING DATE: 2000-09-13  
; NUMBER OF SEQ ID NOS: 75

; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 27  
; LENGTH: 353  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-952-680A-27

Query Match 97.7%; Score 1817; DB 9; Length 353;  
Best Local Similarity 99.2%; Pred. No. 7.5e-146;  
Matches 350; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 MACCLSEBAKARRINDEIRHVRDKRDARRELKLLLGSGSGSTFIKQMRITIHGSG 66  
DB 1 MACCLSEBAKARRINDEIRHVRDKRDARRELKLLLGSGSGSTFIKQMRITIHGSG 60  
QY 67 YSDEDKRGFTKLVYQNIPTAMQAMIRAMDITKIPYKHNKAHQALREVYVERVSAFEN 126  
DB 61 YSDEDKRGFTKLVYQNIPTAMQAMIRAMDITKIPYKHNKAHQALREVYVERVSAFEN 120  
QY 127 PYDAIKSLMNDPGIOECYDRRREYQLSDSTKYYLNDLDRVADPAVLPQODVLRVAVPT 186  
DB 121 PYDAIKSLMNDPGIOECYDRRREYQLSDSTKYYLNDLDRVADPAVLPQODVLRVAVPT 180  
QY 187 TGIIEYFPDLOSIFRNVAVDGGORSEKRMHCEENVTSMFLVALSEYDQVLYESDNEN 246  
DB 181 TGIIEYFPDLOSIFRNVAVDGGORSEKRMHCEENVTSMFLVALSEYDQVLYESDNEN 240  
QY 247 RMEESKALFTIITYPMFONSSVLLFLNKKDLLEKIMYSHLVDFPEYDGPORDAQAAR 306  
DB 241 RMEESKALFTIITYPMFONSSVLLFLNKKDLLEKIMYSHLVDFPEYDGPORDAQAAR 300  
QY 307 EFLIKMFVDLNPDSDKIIYSHFTCATDTENIRFVAAVKDTIQLNLKEYNLV 359  
DB 301 EFLIKMFVDLNPDSDKIIYSHFTCATDTENIRFVAAVKDTIQLNLKEYNLV 353

## RESULT 15

US-09-984-292-16

; Sequence 16, Application US/09984292  
; Patent No. US20020128433A1  
; GENERAL INFORMATION:  
; APPLICANT: YAO, YONG  
; APPLICANT: XU, HONG  
; TITLE OF INVENTION: G-ALPHA-O PROTEIN VARIANTS AND THEIR USE IN THE  
; TITLE OF INVENTION: ANALYSIS AND DISCOVERY OF AGONISTS AND ANTAGONISTS OF  
; FILE REFERENCE: 078003-0280649  
; CURRENT APPLICATION NUMBER: US/09/984,292  
; PRIOR FILING DATE: 2001-10-29  
; PRIOR APPLICATION NUMBER: 60/243,770  
; PRIOR FILING DATE: 2000-10-30  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 16  
; LENGTH: 353  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-984-292-16

Query Match 97.7%; Score 1817; DB 10; Length 353;  
Best Local Similarity 99.2%; Pred. No. 7.5e-146;  
Matches 350; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 MACCLSEBAKARRINDEIRHVRDKRDARRELKLLLGSGSGSTFIKQMRITIHGSG 66  
DB 1 MACCLSEBAKARRINDEIRHVRDKRDARRELKLLLGSGSGSTFIKQMRITIHGSG 60  
QY 67 YSDEDKRGFTKLVYQNIPTAMQAMIRAMDITKIPYKHNKAHQALREVYVERVSAFEN 126  
DB 61 YSDEDKRGFTKLVYQNIPTAMQAMIRAMDITKIPYKHNKAHQALREVYVERVSAFEN 120  
QY 127 PYDAIKSLMNDPGIOECYDRRREYQLSDSTKYYLNDLDRVADPAVLPQODVLRVAVPT 186  
DB 121 PYDAIKSLMNDPGIOECYDRRREYQLSDSTKYYLNDLDRVADPAVLPQODVLRVAVPT 186

Db	121	PYVDAIKSLMNDPGIOECYDRRREYQLSDSTKYIINDLDRAVPAYLPTQODVLRVRPT	180
OY	187	TCIIIEPFDLOSIVFRMVDVGGRSERRKWHCFENVTSIMFLVALSEYDOYLVESDNE	246
Db	181	TCIIIEPFDLOSIVFRMVDVGGRSERRKWHCFENVTSIMFLVALSEYDOYLVESDNE	240
OY	247	RMEESKALERTIIITYPWFONSSVILELNKDLLEEKIMYSHLVDFPEYDGPORDAQAAR	306
Db	241	RMEESKALERTIIITYPWFONSSVILELNKDLLEEKIMYSHLVDFPEYDGPORDAQAAR	300
OY	307	EFILMFVDLNPDSKIIISHFTCATDTENIRFVFAVNDTIIQLNLKEYNLV	359
Db	301	EFILMFVDLNPDSKIIISHFTCATDTENIRFVFAVNDTIIQLNLKEYNAV	353

Search completed: June 30, 2003, 16:19:16  
Job time : 25 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2003, 16:14:36 ; Search time 19 seconds  
(without alignments) 1816.435 Million cell updates/sec

Title: US-09-899-295-2

Perfect score: 1860

Sequence: 1 MTELSIMACCLSEAKFEARR.....VEAAVKDTIIQLNLKEYNLV 359

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: PIR\_73:\*  
2: pir1:\*  
3: pir2:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1857	99.8	359	1 RGM50	GTP-binding regula
2	1844	99.1	359	2 S71963	GTP-binding protei
3	1799	96.7	359	2 S45699	GTP-binding regula
4	1709	91.9	359	2 S30359	GTP-binding regula
5	1685	90.6	359	2 S45700	G-alpha-11 protein
6	1682	90.4	359	1 RGM511	GTP-binding regula
7	1681	90.4	359	1 RGM511	GTP-binding regula
8	1667	89.6	353	2 BA0891	GTP-binding protei
9	1538	82.7	355	2 A40891	GTP-binding protei
10	1535	82.5	355	2 A41534	GTP-binding protei
11	1524	81.9	353	2 S34347	GTP-binding regula
12	1513	81.3	355	2 T15288	hypothetical prote
13	1469.5	79.0	354	2 S33309	GTP-binding regula
14	1425.5	76.6	360	2 JN0115	GTP-binding regula
15	1066	57.3	374	2 BA1534	GTP-binding protei
16	1044	56.1	374	2 A41096	GTP-binding regula
17	913	49.1	353	2 S71964	GTP-binding protei
18	904	48.6	354	1 RGHU11	GTP-binding regula
19	898.5	48.3	354	1 RGHU11	GTP-binding regula
20	898.5	48.3	354	1 RGHU11	GTP-binding regula
21	897.5	48.3	354	1 RGHU11	GTP-binding regula
22	895.5	48.1	354	1 S40508	GTP-binding regula
23	893	48.0	353	2 JC7661	G protein alpha su
24	890	47.8	353	2 T50482	GTP-binding regula
25	888.5	47.8	354	1 RGLX11	GTP-binding regula
26	888.5	47.8	354	2 I50237	GTP-binding regula
27	887.5	47.7	354	2 S40509	GTP-binding regula
28	886.5	47.7	354	1 RGHU02	GTP-binding regula
29	886.5	47.7	354	2 S28157	GTP-binding regula

30	883	47.5	355	2 S28158	GTP-binding regula
31	882.5	47.4	354	1 RGHU13	GTP-binding regula
32	882.5	47.4	354	1 RGHU13	GTP-binding regula
33	882	47.4	355	2 I50238	G12 protein alpha
34	881.5	47.4	354	1 RGHU13	GTP-binding regula
35	878.5	47.2	354	2 S27014	GTP-binding regula
36	877.5	47.2	354	2 S24362	GTP-binding regula
37	877	47.2	355	1 RGHU12	GTP-binding regula
38	877	47.2	355	2 A61031	GTP-binding regula
39	876.5	47.1	354	2 S28159	GTP-binding regula
40	873.5	47.0	354	2 S27013	GTP-binding regula
41	872.5	46.9	354	1 RGM502	GTP-binding regula
42	872.5	46.9	354	1 RGHU02	GTP-binding regula
43	871.5	46.9	354	1 RGHU02	GTP-binding regula
44	871	46.8	355	1 RGHU02	GTP-binding regula
45	865.5	46.5	354	1 RGHU01	GTP-binding regula

ALIGNMENTS

RESULT 1

RGM50

GTP-binding regulatory protein Gq alpha chain - mouse

N:Alternate names: guanine nucleotide binding protein Gq alpha chain; heterotrimeric

C:Species: Mus musculus (house mouse)

C:Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 19-Jan-2001

C:Accession: A38414

R:Stratmann, M.; Simon, M.I.

A:Title: G protein diversity: a distinct class of alpha subunits is present in verte

A:Reference number: A38414; MUID:91067657; PMID:2123549

A:Accession: A38414

A:Molecule type: mRNA

A:Residues: 1-359 <STR>

A:Cross-References: GB:M55412; NID:9193501; PIDN:AAA63306.1; PID:9193502

C:Comment: The G proteins are a family of guanine nucleotide-binding proteins that re

ases. The beta and gamma chains, required for GTPase activity, appear to be common to

raise: it is specific for each type of G protein.

C:Superfamily: GTP-binding regulatory protein Gs alpha chain

C:Keywords: GTP binding; heterotrimer; nucleotide binding; P-loop; signal transductio

F:46-53/Region: nucleotide binding motif A (P-loop)

F:274-277/Region: GTP-binding NKXD motif

F:52/Binding site: GTP (Lys) #status predicted

F:183/Modified site: ADP-riboylarginine (Arg) (by cholera toxin) #status predicted

Query Match

99.8%; Score 1857; DB 1; Length 359;

Best Local Similarity 99.7%; Pred. No. 1.4e-126;

Matches 358; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY	1	MTELSIMACCLSEAKFEARRINDEIRHRRKRDARRELKLLGTGSGSKSTFKOMR	60
DB	1	MTELSIMACCLSEAKFEARRINDEIRHRRKRDARRELKLLGTGSGSKSTFKOMR	60
QY	61	IIHSGSYDEDEKRGFTKLYONIIFAMQAMIRAMPDLKIPYKEHNKAHAQVREVDYK	120
DB	61	IIHSGSYDEDEKRGFTKLYONIIFAMQAMIRAMPDLKIPYKEHNKAHAQVREVDYK	120
QY	121	VSAFENPYDAIKSLMNDPGIOECYDRRREYLOSLSTKYLLNDLRVADPAVLPOQDYL	180
DB	121	VSAFENPYDAIKSLMNDPGIOECYDRRREYLOSLSTKYLLNDLRVADPAVLPOQDYL	180
QY	181	RVRVPTGTITIEYFPDQSYIFRMVVGGRSRKRWIHCFENVTSIMFVALSEYDQYLV	240
DB	181	RVRVPTGTITIEYFPDQSYIFRMVVGGRSRKRWIHCFENVTSIMFVALSEYDQYLV	240
QY	241	ESDNNRMEESKALFRTITTYPPFONSSVILFLNKKDLLEKIMTSHLDVPEYDQGR	300
DB	241	ESDNNRMEESKALFRTITTYPPFONSSVILFLNKKDLLEKIMTSHLDVPEYDQGR	300
QY	301	DAQAAREFLKMFVDLNPDSKIIYSHFTCATDTEINIRVFAAVKDTIIQLNLKEYNLV	359
DB	301	DAQAAREFLKMFVDLNPDSKIIYSHFTCATDTEINIRVFAAVKDTIIQLNLKEYNLV	359





```

Db 181 RVRPTTGIIIEYPPDLNIIIFRMVDVGORSEKRWIHCEFNWTSIMFLVALSEYDQVL 240
QY 241 ESDNENMEESKALFRITITYPWFONSSVILFLNKKDLLEKIMYSHLVDFPEYDQVR 300
Db 241 ESDNENMEESKALFRITITYPWFONSSVILFLNKKDLLEKIMYSHLVDFPEYDQVR 300
QY 301 DAQAAREFIILKMFVDLNPDSDKITYSHTCATDTEENIRFVFAAVKDTILQNLKEYNLV 359
Db 301 DAQAAREFIILKMFVDLNPDSDKITYSHTCATDTEENIRFVFAAVKDTILQNLKEYNLV 359

```

## RESULT 5

```

S45700
G-alpha-11 protein - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C>Date: 10-Dec-1994 #sequence_revision 10-Nov-1995 #text_change 02-Feb-2001
C:Accession: S45700
R:Shapira, H.; Way, J.; Lipinsky, D.; Oron, Y.; Battey, J.F.
FEBS Lett. 348, 89-92, 1994
A:Title: Neuromedin B receptor, expressed in Xenopus laevis oocytes, selectively couples
A:Reference number: S45699; MUID:94298961; PMID:8026589
A:Accession: S45700
A:Molecule type: mRNA
A:Residues: 1-359 <SHA>
A:Cross-references: GB:U10494; NID:9505687; PIDN:AAA52188.1; PID:9505688
C:Superfamily: GTP-binding regulatory protein Gs alpha chain
C:Keywords: GTP binding; nucleotide binding; P-loop
F:46-53/Region: nucleotide-binding motif A (P-loop)
F:274-277/Region: GTP-binding NKXD motif

```

```

Query Match 90.6%; Score 1685; DB 2; Length 359;
Best Local Similarity 88.9%; Pred. No. 3.5e-114;
Matches 319; Conservative 25; Mismatches 15; Indels 0; Gaps 0;

```

```

QY 1 MTELSIMACCLSEAKERRINDIEKRRDKRDARRELKLLLGESGSKSTFIKQMR 60
Db 1 MTELSIMACCLSEAKERRINDIEKRRDKRDARRELKLLLGESGSKSTFIKQMR 60
QY 61 ITHSGYSDEDKRGFTLVQNIPTAMQAMRAMDTLKIPKYEHKNAHQLVREYVEK 120
Db 61 ITHSGYSDEDKRGFTLVQNIPTAMQAMRAMDTLKIPKYEHKNAHQLVREYVEK 120
QY 121 VSAFENPYDAIKSLMNDPGIOECYDRRREYQSDSTKYYLINDLRVADPAVLPDQDVL 180
Db 121 VSAFENPYDAIKSLMNDPGIOECYDRRREYQSDSTKYYLINDLRVADPAVLPDQDVL 180
QY 181 RVRPTTGIIIEYPPDLNIIIFRMVDVGORSEKRWIHCEFNWTSIMFLVALSEYDQVL 240
Db 181 RVRPTTGIIIEYPPDLNIIIFRMVDVGORSEKRWIHCEFNWTSIMFLVALSEYDQVL 240
QY 241 ESDNENMEESKALFRITITYPWFONSSVILFLNKKDLLEKIMYSHLVDFPEYDQVR 300
Db 241 ESDNENMEESKALFRITITYPWFONSSVILFLNKKDLLEKIMYSHLVDFPEYDQVR 300
QY 301 DAQAAREFIILKMFVDLNPDSDKITYSHTCATDTEENIRFVFAAVKDTILQNLKEYNLV 359
Db 301 DAQAAREFIILKMFVDLNPDSDKITYSHTCATDTEENIRFVFAAVKDTILQNLKEYNLV 359

```

## RESULT 6

```

RGMS11
GTP-binding regulatory protein G11 alpha chain - mouse
N:Alternate names: guanine nucleotide binding protein G11 alpha chain; heterotrimeric G-
C:Species: Mus musculus (house mouse)
C>Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001
C:Accession: B38414
R:Strathmann, M.; Simon, M.I.
Proc. Natl. Acad. Sci. U.S.A. 87, 9113-9117, 1990
A:Title: G protein diversity: a distinct class of alpha subunits is present in vertebrate
A:Reference number: A38414; MUID:91067657; PMID:2123549
A:Molecule type: mRNA

```

```

A:Residues: 1-359 <STR>
A:Cross-references: GB:M55411; NID:q193499; PIDN:AAA63305.1; PID:q193500
C:Comment: The G proteins are a family of guanine nucleotide-binding proteins that re-
ains. The beta and gamma chains, required for GTPase activity, appear to be common to
rase. It is specific for each type of G protein.
C:Superfamily: GTP-binding regulatory protein Gs alpha chain
C:Keywords: GTP binding; heterotrimer; nucleotide binding; P-loop; signal transductio
F:46-53/Region: nucleotide-binding motif A (P-loop)
F:274-277/Region: GTP-binding NKXD motif
F:52/Binding site: GTP (lys) #status predicted
F:183/Modified site: ADP-ribosylarginine (Arg) (by cholera toxin) #status predicted

```

```

Query Match 90.4%; Score 1682; DB 1; Length 359;
Best Local Similarity 88.6%; Pred. No. 5.7e-114;
Matches 318; Conservative 28; Mismatches 13; Indels 0; Gaps 0;

```

```

QY 1 MTELSIMACCLSEAKERRINDIEKRRDKRDARRELKLLLGESGSKSTFIKQMR 60
Db 1 MTELSIMACCLSEAKERRINDIEKRRDKRDARRELKLLLGESGSKSTFIKQMR 60
QY 61 ITHSGYSDEDKRGFTLVQNIPTAMQAMRAMDTLKIPKYEHKNAHQLVREYVEK 120
Db 61 ITHSGYSDEDKRGFTLVQNIPTAMQAMRAMDTLKIPKYEHKNAHQLVREYVEK 120
QY 121 VSAFENPYDAIKSLMNDPGIOECYDRRREYQSDSTKYYLINDLRVADPAVLPDQDVL 180
Db 121 VSAFENPYDAIKSLMNDPGIOECYDRRREYQSDSTKYYLINDLRVADPAVLPDQDVL 180
QY 181 RVRPTTGIIIEYPPDLNIIIFRMVDVGORSEKRWIHCEFNWTSIMFLVALSEYDQVL 240
Db 181 RVRPTTGIIIEYPPDLNIIIFRMVDVGORSEKRWIHCEFNWTSIMFLVALSEYDQVL 240
QY 241 ESDNENMEESKALFRITITYPWFONSSVILFLNKKDLLEKIMYSHLVDFPEYDQVR 300
Db 241 ESDNENMEESKALFRITITYPWFONSSVILFLNKKDLLEKIMYSHLVDFPEYDQVR 300
QY 301 DAQAAREFIILKMFVDLNPDSDKITYSHTCATDTEENIRFVFAAVKDTILQNLKEYNLV 359
Db 301 DAQAAREFIILKMFVDLNPDSDKITYSHTCATDTEENIRFVFAAVKDTILQNLKEYNLV 359

```

## RESULT 7

```

RGHDY
GTP-binding regulatory protein Gy alpha chain - human
N:Alternate names: guanine nucleotide binding protein Gy alpha chain; heterotrimeric
C:Species: Homo sapiens (man)
C>Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001
C:Accession: A39394
R:Jiang, M.; Pandey, S.; Tran, V.T.; Fong, H.K.W.
Proc. Natl. Acad. Sci. U.S.A. 88, 3907-3911, 1991
A:Title: Guanine nucleotide-binding regulatory proteins in retinal pigment epithelial
A:Reference number: A39394; MUID:91219481; PMID:1902575
A:Accession: A39394
A:Molecule type: mRNA
A:Residues: 1-359 <JTA>
A:Cross-references: GB:M69013; NID:g183690; PIDN:AAA58624.1; PID:g183691
C:Comment: The G proteins are a family of guanine nucleotide-binding proteins that re-
ains. The beta and gamma chains, required for GTPase activity, appear to be common to
rase. It is specific for each type of G protein.
C:Genetics:
A:Gene: GDB:GNAL1
A:Map position: 19p13.3-19p13.3
A:Superfamily: GTP-binding regulatory protein Gs alpha chain
C:Keywords: GTP binding; heterotrimer; nucleotide binding; P-loop; signal transductio
F:46-53/Region: nucleotide-binding motif A (P-loop)
F:274-277/Region: GTP-binding NKXD motif
F:52/Binding site: GTP (lys) #status predicted
F:183/Modified site: ADP-ribosylarginine (Arg) (by cholera toxin) #status predicted

```

```

Query Match 90.4%; Score 1681; DB 1; Length 359;
Best Local Similarity 89.1%; Pred. No. 6.7e-114;
Matches 320; Conservative 23; Mismatches 16; Indels 0; Gaps 0;

```



A:Residues: 217-267 <STR>  
A:Cross-references: GB:M57616; NID:g193380; PIDN:AAA6304.1; PID:g193381; GB:M26739  
C:Superfamily: GTP-binding regulatory protein Gs alpha chain  
C:Keywords: GTP binding; nucleotide binding; P-loop  
F:42-49/Region: nucleotide-binding motif A (P-loop)  
F:153-154/Region: GTP-binding motif A (P-loop)  
F:153-154/Region: GTP-binding motif A (P-loop)  
F:270-273/Region: GTP-binding motif A (P-loop)

Query Match 82.5%; Score 1535; DB 2; Length 355;  
Best Local Similarity 82.1%; Pred. No. 1.4e-103;  
Matches 288; Conservative 35; Mismatches 28; Indels 0; Gaps 0;

QY 9 CCLSEAEKARRINDEIERHVRDKRDARRELKLLLTGSGSKSTFIKQRIIHGSGS 68  
DB 5 CCLSAEKEQSRISEIERHVRDKRDARRELKLLLTGSGSKSTFIKQRIIHGSGS 64  
QY 69 DEDKRGFTLVYONIFTMQAMIRAMDITKIPYKEHNKAQAOLVREVDKESAFENPY 128  
DB 65 DEDKRGFTLVYONIFTMQAMIRAMDITKIPYKEHNKAQAOLVREVDKESAFENPY 124  
QY 129 VDAIKSLMNDPGIOECYDRREYOLSDSTKYIYNDLRVADPAVLPQODVLRVAVPTTG 188  
DB 125 VAAIKQMLWDPGIOECYDRREYOLSDSAKYIYNDLRVADPAVLPQODVLRVAVPTTG 184  
QY 189 IIEYFPDQSVIFRMVDVGQSRERRKWHCFENVTSMFLVALSEYDQVLVESDNENRM 248  
DB 185 IIEYFPDQSVIFRMVDVGQSRERRKWHCFENVTSMFLVALSEYDQVLVESDNENRM 244  
QY 249 EESKALPFTITTYPMFQSSVILFLNKKDLLEKIMSHLYDPEYDQVLRVAVPTTG 308  
DB 245 EESKALPFTITTYPMFQSSVILFLNKKDLLEKIMSHLYDPEYDQVLRVAVPTTG 304  
QY 309 ILKKEFVDLPDSDKIIYSHFTCATDTENIRFVAFAVKTIIQLNKEYNLV 359  
DB 305 ILKKEFVDLPDSDKIIYSHFTCATDTENIRFVAFAVKTIIQLNKEYNLV 355

## RESULT 11

S34347  
GTP-binding regulatory protein Gq alpha chain - great pond snail  
C:Species: Lymnaea stagnalis (great pond snail)  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 02-Feb-2001  
C:Accession: S65461; S34347  
R:KNO1, J.C.; Ramatstingh, S.; van Kesteren, E.R.; van Minnen, J.; Plantia, R.J.; van Hee  
Eur. J. Biochem. 230, 193-199, 1995  
A:Title: Cloning of a molluscan G protein alpha subunit of the Gq class which is expressed  
A:Reference number: S65461; MUID:95324523; PMID:7601100  
A:Accession: S65461  
A:Molecule type: mRNA

A:Residues: 1-353 <KNO>  
A:Cross-references: EMBL:Z23106; NID:g312629; PIDN:CAA80653.1; PID:g312630  
C:Superfamily: GTP-binding regulatory protein Gs alpha chain  
C:Keywords: GTP binding; heterotrimer; nucleotide binding; P-loop; signal transduction  
F:40-48/Region: nucleotide-binding motif A (P-loop)  
F:150-152/Region: GTP-binding motif A (P-loop)  
F:268-271/Region: GTP-binding motif A (P-loop)  
F:46/Binding site: GTP (Lys) #status predicted  
F:177/Modified site: ADP-ribosearginine (Arg) (by cholera toxin) #status predicted

Query Match 81.9%; Score 1524; DB 2; Length 353;  
Best Local Similarity 81.6%; Pred. No. 1.4e-102;  
Matches 288; Conservative 29; Mismatches 36; Indels 0; Gaps 0;

QY 7 MACCLSEAEKARRINDEIERHVRDKRDARRELKLLLTGSGSKSTFIKQRIIHGSG 66  
DB 1 MACCLSEAEKARRINDEIERHVRDKRDARRELKLLLTGSGSKSTFIKQRIIHGSG 60  
QY 67 YSDEKRGFTLVYONIFTMQAMIRAMDITKIPYKEHNKAQAOLVREVDKESAFEN 126  
DB 61 YSDEKRGFTLVYONIFTMQAMIRAMDITKIPYKEHNKAQAOLVREVDKESAFEN 120  
QY 127 PYVAIKSLMNDPGIOECYDRREYOLSDSTKYIYNDLRVADPAVLPQODVLRVAVPTTG 186  
DB 127 PYVAIKSLMNDPGIOECYDRREYOLSDSTKYIYNDLRVADPAVLPQODVLRVAVPTTG 186

DB 121 PCVDAIISLMDNDGIOECYDRREYOLSDSAKYIYNDLRVADPAVLPQODVLRVAVPTTG 180  
QY 187 TGIIEYFPDQSVIFRMVDVGQSRERRKWHCFENVTSMFLVALSEYDQVLVESDNEN 246  
DB 181 TGIIEYFPDQSVIFRMVDVGQSRERRKWHCFENVTSMFLVALSEYDQVLVESDNEN 240  
QY 247 RMEESKALPFTITTYPMFQSSVILFLNKKDLLEKIMSHLYDPEYDQVLRVAVPTTG 306  
DB 241 RMEESKALPFTITTYPMFQSSVILFLNKKDLLEKIMSHLYDPEYDQVLRVAVPTTG 300  
QY 307 EFILKMEVLDLPDSDKIIYSHFTCATDTENIRFVAFAVKTIIQLNKEYNLV 359  
DB 301 EFILKMEVLDLPDSDKIIYSHFTCATDTENIRFVAFAVKTIIQLNKEYNLV 353

## RESULT 12

T15288  
hypothetical protein M01D7.7 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 24-Nov-1999  
C:Accession: T15288

R:Gartung, S.; Goela, D.; Wilson, R.  
submitted to the EMBL Data Library, May 1997

A:Description: The sequence of C. elegans cosmid M01D7.

A:Reference number: Z18322

A:Accession: T15288

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-355 <GAT>

A:Cross-references: EMBL:AF003739; NID:g2105482; PID:g2105489; PIDN:AAB58071.1; GSPDB

A:Experimental source: strain Bristol N2; clone M01D7

C:Genetics:

A:Gene: CESP:M01D7.7

A:Map position: 1

A:Introns: 40/1; 82/3; 155/2; 198/2; 241/3; 293/1; 329/1

C:Superfamily: GTP-binding regulatory protein Gs alpha chain

Query Match 81.3%; Score 1513; DB 2; Length 355;  
Best Local Similarity 82.0%; Pred. No. 8.7e-102;  
Matches 291; Conservative 26; Mismatches 36; Indels 2; Gaps 1;

QY 7 MACCLSEAEKARRINDEIERHVRDKRDARRELKLLLTGSGSKSTFIKQRIIHGSG 66  
DB 1 MACCLSEAEKARRINDEIERHVRDKRDARRELKLLLTGSGSKSTFIKQRIIHGSG 60  
QY 67 YSDEKRGFTLVYONIFTMQAMIRAMDITKIPYKEHNKAQAOLVREVDKESAFEN 124  
DB 61 YSDEKRGFTLVYONIFTMQAMIRAMDITKIPYKEHNKAQAOLVREVDKESAFEN 120  
QY 125 ENPYVDAIKSLMNDPGIOECYDRREYOLSDSTKYIYNDLRVADPAVLPQODVLRVAVPTTG 184  
DB 121 EEPYVSYKELMEDSGIOECYDRREYOLSDSAKYIYNDLRVADPAVLPQODVLRVAVPTTG 180  
QY 185 PTTGIIIEYFPDQSVIFRMVDVGQSRERRKWHCFENVTSMFLVALSEYDQVLVESDN 244  
DB 181 PTTGIIIEYFPDQSVIFRMVDVGQSRERRKWHCFENVTSMFLVALSEYDQVLVESDN 240  
QY 245 ENRMEESKALPFTITTYPMFQSSVILFLNKKDLLEKIMSHLYDPEYDQVLRVAVPTTG 304  
DB 241 ENRMEESKALPFTITTYPMFQSSVILFLNKKDLLEKIMSHLYDPEYDQVLRVAVPTTG 300  
QY 305 AREFLKMEVLDLPDSDKIIYSHFTCATDTENIRFVAFAVKTIIQLNKEYNLV 359  
DB 301 AREFLKMEVLDLPDSDKIIYSHFTCATDTENIRFVAFAVKTIIQLNKEYNLV 355

## RESULT 13

S33309  
GTP-binding regulatory protein Gq alpha chain - northern European squid

C:Species: Loligo forbesi (northern European squid)

C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 02-Feb-2001

C:Accession: S33309

R:Kyda, N.D.P.; Findlay, J.B.C.; Reid, J.D.

Biochem. J. 292, 333-341, 1993  
 A:Title: The molecular cloning of the squid (*Loligo forbesi*) visual Gq-alpha subunit and  
 A:Reference number: S33309; MUID:93277493; PMID:8503868  
 A:Accession: S33309  
 A:Status: Preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-354 <RVD>  
 A:Cross-references: EMBL:L10289  
 C:Superfamily: GTP-binding regulatory protein Gs alpha chain  
 C:Keywords: GTP binding; nucleotide binding; P-loop  
 F:40-47/Region: nucleotide-binding motif A (P-loop)  
 F:150-152/Region: GTP-binding SAK/L motif  
 F:269-272/Region: GTP-binding NKXD motif

Query Match 79.0%; Score 1469.5; DB 2; Length 354;  
 Best Local Similarity 77.4%; Pred. No. 1.2e-98;  
 Matches 274; Conservative 39; Mismatches 40; Indels 1; Gaps 1;

QY 7 MACCLSEAKARRINDEIERHVRDKRDARRELKLLLGESGSKSTFIQOMRIHSG 66  
 DB 1 MACCLSEAKARRINDEIERHVRDKRDARRELKLLLGESGSKSTFIQOMRIHSG 60  
 QY 67 YSDEKRGFTKLVQNIPTAMQAMIRAMDTLKIPYKEHNKAHQAOLVREVEKVSAPEN 126  
 DB 61 YSEDRKGFTKLVQNIPTAMQAMIRAMDTLKIPYKEHNKAHQAOLVREVEKVSAPEN 120  
 QY 127 PYVDATKSLMNDPGIOECYDRRREYQLSDSTKYLLNDLRVADPAYLPTQDVLRAVRPT 186  
 DB 121 GHVTAIKGWTDPGMEQCYDRRREYQLSDSTKYLLNDLRVADPAYLPTQDVLRAVRPT 180  
 QY 187 TGIIEFPDLOSIVFRWVDGQSRERKRWIHCENVTSMFLVALSEYDOVLESNDN-E 245  
 DB 181 TGIIEFPDLOSIVFRWVDGQSRERKRWIHCENVTSMFLVALSEYDOVLESNDN-E 240  
 QY 246 NRMESEKALFRTITTYPMFQNSVILFLNKKDLLEKIMSHLVDPEDYDGPORDAQA 305  
 DB 241 NRMESEKALFRTITTYPMFQNSVILFLNKKDLLEKIMSHLVDPEDYDGPORDAQA 300  
 QY 306 REFLLKMFVDLNPDSKTIYSHFTCATDTENIRFVAVKDTIILNLKEYNLV 359  
 DB 301 REFLLKMFVDLNPDSKTIYSHFTCATDTENIRFVAVKDTIILNLKEYNLV 354

## RESULT 14

JN0115  
 GTP-binding regulatory protein dqg alpha chain - fruit fly (*Drosophila melanogaster*)  
 N:Alternate names: dqg protein  
 C:Species: *Drosophila melanogaster*  
 C:Date: 12-Feb-1993 #sequence\_revision 12-Feb-1993 #text\_change 02-Feb-2001  
 C:Accession: JN0115  
 R:Lee, Y.-J.; Dobbs, M.B.; Verardi, M.L.; Hyde, D.R.  
 Neuron 5, 889-898, 1990  
 A:Title: dqg: a *Drosophila* gene encoding a visual system-specific G alpha molecule.  
 A:Reference number: JN0115; MUID:91097801; PMID:2125225  
 A:Accession: JN0115  
 A:Molecule type: DNA  
 A:Residues: 1-360 <LEE>  
 C:Genetics:  
 A:Gene: dqg  
 A:Cross-references: Flybase:FBgn0004435  
 A:Introns: 40/1; 101/3; 153/2; 196/2; 239/3; 298/1; 334/1  
 C:Superfamily: GTP-binding regulatory protein Gs alpha chain  
 C:Keywords: alternative splicing; GTP binding; nucleotide binding; P-loop  
 F:1-360/Product: GTP-binding regulatory protein dqg alpha chain I #status predicted <MAT  
 F:1-297,333-360/Product: GTP-binding regulatory protein dqg alpha chain II #status predi  
 F:40-47/Region: nucleotide-binding motif A (P-loop)  
 F:150-152/Region: GTP-binding SAK/L motif  
 F:268-271/Region: GTP-binding NKXD motif

Query Match 76.6%; Score 1425.5; DB 2; Length 360;  
 Best Local Similarity 75.2%; Pred. No. 1.8e-95;  
 Matches 270; Conservative 37; Mismatches 45; Indels 7; Gaps 1;

QY 7 MACCLSEAKARRINDEIERHVRDKRDARRELKLLLGESGSKSTFIQOMRIHSG 66  
 DB 1 MACCLSEAKARRINDEIERHVRDKRDARRELKLLLGESGSKSTFIQOMRIHSG 60  
 QY 67 YSDEKRGFTKLVQNIPTAMQAMIRAMDTLKIPYKEHNKAHQAOLVREVEKVSAPEN 126  
 DB 61 YSEDRKGFTKLVQNIPTAMQAMIRAMDTLKIPYKEHNKAHQAOLVREVEKVSAPEN 120  
 QY 127 PYVDATKSLMNDPGIOECYDRRREYQLSDSTKYLLNDLRVADPAYLPTQDVLRAVRPT 186  
 DB 121 GHVTAIKGWTDPGMEQCYDRRREYQLSDSTKYLLNDLRVADPAYLPTQDVLRAVRPT 180  
 QY 187 TGIIEFPDLOSIVFRWVDGQSRERKRWIHCENVTSMFLVALSEYDOVLESNDN 246  
 DB 181 TGIIEFPDLOSIVFRWVDGQSRERKRWIHCENVTSMFLVALSEYDOVLESNDN 240  
 QY 247 NRMESEKALFRTITTYPMFQNSVILFLNKKDLLEKIMSHLVDPEDYDGPORDAQA 300  
 DB 241 NRMESEKALFRTITTYPMFQNSVILFLNKKDLLEKIMSHLVDPEDYDGPORDAQA 300  
 QY 300 RDAQAAREFLKMFVDLNPDSKTIYSHFTCATDTENIRFVAVKDTIILNLKEYNLV 358  
 DB 301 RDAQAAREFLKMFVDLNPDSKTIYSHFTCATDTENIRFVAVKDTIILNLKEYNLV 359

## RESULT 15

B41534  
 GTP-binding protein alpha-15 chain - mouse  
 C:Species: *Mus musculus* (house mouse)  
 C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 02-Feb-2001  
 R:Wilkie, T.M.; Scherle, P.A.; Strathmann, M.P.; Slepak, V.Z.; Simon, M.I.  
 Proc. Natl. Acad. Sci. U.S.A. 89, 10049-10053, 1991  
 A:Title: Characterization of G-protein alpha subunits in the G-q class: expression in  
 A:Reference number: B41534; MUID:92052208; PMID:1946421  
 A:Accession: B41534  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-374 <WIL>  
 A:Cross-references: GB:M0632; NID:g193570; PIDN:AAA37713.1; PID:g193571  
 C:Superfamily: GTP-binding regulatory protein Gs alpha chain  
 C:Keywords: GTP binding; nucleotide binding; P-loop  
 F:49-56/Region: nucleotide-binding motif A (P-loop)  
 F:277-280/Region: GTP-binding NKXD motif

Query Match 57.3%; Score 1066; DB 2; Length 374;  
 Best Local Similarity 57.2%; Pred. No. 1.6e-69;  
 Matches 207; Conservative 55; Mismatches 88; Indels 12; Gaps 2;

QY 10 CLSEAKARRINDEIERHVRDKRDARRELKLLLGESGSKSTFIQOMRIHSGYSD 69  
 DB 13 CLSEAKARRINDEIERHVRDKRDARRELKLLLGESGSKSTFIQOMRIHSGYSD 69  
 QY 70 EDRKRGFTKLVQNIPTAMQAMIRAMDTLKIPYKEHNKAHQAOLVREVEKVSAPENPV 129  
 DB 73 EDRKRGFTKLVQNIPTAMQAMIRAMDTLKIPYKEHNKAHQAOLVREVEKVSAPENPV 132  
 QY 130 DAISLNDNDPGIOECYDRRREYQLSDSTKYLLNDLRVADPAYLPTQDVLRAVRPTGI 189  
 DB 133 VAMQYILWMDAGIACRYERRRREYQLSDSTKYLLNDLRVADPAYLPTQDVLRAVRPTGI 192  
 QY 190 IEYFPDLOSIVFRWVDGQSRERKRWIHCENVTSMFLVALSEYDOVLESNDN 249  
 DB 193 IEYFPDLOSIVFRWVDGQSRERKRWIHCENVTSMFLVALSEYDOVLESNDN 252  
 QY 250 ESKALFRTITTYPMFQNSVILFLNKKDLLEKIMSHLVDPEDYDGPORDAQAAREFI 309  
 DB 253 ESKALFRTITTYPMFQNSVILFLNKKDLLEKIMSHLVDPEDYDGPORDAQAAREFI 312  
 QY 310 LKMFVDL-----NPDS-----KTIYSHFTCATDTENIRFVAVKDTIILNLKEYN 357  
 DB 313 LKMFVDL-----NPDS-----KTIYSHFTCATDTENIRFVAVKDTIILNLKEYN 372

Tue Jul 1 09:54:47 2003

us-09-899-295-2.rpr

Page 7

Oy	358	LV	359
		I :	
Db	373	LL	374

Search completed: June 30, 2003, 16:18:23  
Job time : 20 secs

---



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2003, 16:10:20 ; Search time 11 Seconds  
(without alignments)  
1353.637 Million cell updates/sec

Title: US-09-899-295-2  
Perfect score: 1860  
Sequence: 1 MTELSIMACLSSEAKEARR.....VFAVKDTILQLNLKEYNLV 359

Scoring table: BLOSUM62  
Gapop 10.0 , Gapect 0.5

Searched: 112892 seqs, 41476328 residues  
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	1830	98.4	353 1	GBQ_MOUSE
2	1822	98.0	353 1	GBQ_CANFA
3	1818	97.7	353 1	GBQ_RAT
4	1817	97.7	353 1	GBQ_HUMAN
5	1766	94.9	353 1	GBQ_XENLA
6	1709	91.9	359 1	GB11_MELGA
7	1691	90.9	359 1	GB11_BOVIN
8	1690	90.9	359 1	GB11_HUMAN
9	1685	90.6	359 1	GB11_XENLA
10	1682	90.4	359 1	GB11_MOUSE
11	1674	90.0	359 1	GB11_RAT
12	1538	82.7	355 1	GB14_BOVIN
13	1537	82.6	354 1	GB14_XENLA
14	1535	82.5	355 1	GB14_MOUSE
15	1530	82.3	353 1	GBQ_PATYE
16	1529	82.2	353 1	GBQ_HOMAM
17	1528	82.2	355 1	GB14_HUMAN
18	1524	81.9	353 1	GBQ_LYMET
19	1469.5	79.0	354 1	GBQ_LOURO
20	1439	77.4	353 1	GBQ_DROME
21	1066	57.3	374 1	GB15_MOUSE
22	1052	56.6	374 1	GB15_RAT
23	1044	56.1	374 1	GB15_HUMAN
24	908	48.8	352 1	GB11_EMENT
25	904	48.6	352 1	GB11_CANFA
26	896	48.2	352 1	GB11_COEHE
27	893.5	48.0	353 1	GB11_HUMAN
28	892.5	48.0	353 1	GB11_RAT
29	891	47.9	352 1	GBA1_CRYPA
30	890	47.8	353 1	GBA1_NEUCR
31	889	47.7	352 1	GBA1_SPOSC
32	886.5	47.7	352 1	GB11_CAVPO
33	886	47.6	352 1	GBA1_MAGGR

34	884	47.5	352 1	GBA1_COLTR	042784 collettotic
35	883.5	47.5	353 1	GB11_CHICK	P50146 gallus gall
36	883.5	47.5	353 1	GB11_XENLA	P27044 xenopus lae
37	881.5	47.4	353 1	GB02_HUMAN	P29777 homo sapien
38	878.5	47.2	353 1	GBAK_CAVPO	P38403 cavia porce
39	877.5	47.2	353 1	GB02_CRILO	P17806 cricetus
40	877.5	47.2	353 1	GBAK_RAT	P08753 rattus norv
41	877.5	47.2	353 1	GBH2_MOUSE	P50149 mus musculu
42	877	47.2	354 1	GB12_CHICK	P08774 gallus gall
43	876.5	47.1	353 1	GBAK_HUMAN	P08754 homo sapien
44	874	47.0	354 1	GB12_CAVPO	P38402 cavia porce
45	874	47.0	354 1	GB12_ORVLA	O13055 oryzias lat

ALIGNMENTS

RESULT 1  
GBQ\_MOUSE STANDARD: PRT: 353 AA.  
ID GBQ\_MOUSE  
AC P21279: 01-MAY-1991 (Rel. 18, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Guanine nucleotide-binding protein G(q), alpha subunit.  
GN GNAO.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-Brain;  
RX MEDLINE=94043367; PubMed=2123549;  
RA "Stratmann M., Simon M.I.;  
RT "G protein diversity: a distinct class of alpha subunits is present  
RT in vertebrates and invertebrates."  
RL Proc. Natl. Acad. Sci. U.S.A. 87:9113-9117(1990).  
RN [2]  
RP PALMITOYLATION.  
RC TISSUE-Brain;  
RX MEDLINE=94043367; PubMed=8227063;  
RA "Medeggerter P.B., Chu D.H., Wilson P.T., Lewis M.J., Bourne H.R.;  
RT "Palmitoylation is required for signalling functions and membrane  
RT attachment of Gq alpha and Gs alpha."  
RL J. Biol. Chem. 268:25001-25008(1993).  
CC -1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE  
CC INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE  
CC SIGNALING SYSTEMS.  
CC -1- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).  
CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.  
CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(O)).  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC  
CC EMBL: M55412; AAA63306.1; ALT\_INIT.  
CC PIR: A38414; RGM5O.  
CC HSSP: P10824; IBOF.  
CC MGD: MGI:95776; Gnaq.  
CC InterPro: IPR001019; Gprotein\_alpha.  
CC Pfam: PF00503; G-alpha; 1.  
CC PRINTS: PR00318; GPROTEINA.  
CC PRODOM: PD000281; Gprotein\_alpha; 1.  
CC SMART: SM00275; G-alpha; 1.  
CC GMP-binding; Transducer; Multigene family; ADP-ribosylation;  
CC Palmitate; Lipoprotein; 3  
CC LIPID 3  
CC PALMITATE.

FT LIPID 4 4 PALMITATE.  
 FT NP\_BIND 40 47 GTP (BY SIMILARITY).  
 FT NP\_BIND 199 203 GTP (BY SIMILARITY).  
 FT NP\_BIND 268 271 GTP (BY SIMILARITY).  
 FT MOD\_RES 177 177 ADP-RIBOSYL(1) (BY ACTION OF CTX)  
 (BY SIMILARITY)  
 SQ SEQUENCE 353 AA: 41478 MW: 418261.674 Cys64;

Query Match 98.4%; Score 1830; DB 1; Length 353;  
 Best Local Similarity 99.7%; Pred. No. 2.5e-130;  
 Matches 352; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 MACCLSEAEKARRINDEIRHVRKRDARRELKLLLTGSGSKSTFKOKRIIHSG 66  
 1 MACCLSEAEKARRINDEIRHVRKRDARRELKLLLTGSGSKSTFKOKRIIHSG 60  
 QY YSDEKRGFTKLYONFTMQAMIRAMDTLKIPYKEHNKHAQAOLREVDEKVSFEN 126  
 61 YSDEKRGFTKLYONFTMQAMIRAMDTLKIPYKEHNKHAQAOLREVDEKVSFEN 120  
 Db PYVDATKSLWNDPGIOECYDRREYOLSDSTKYLLDLDVADPAVLPQODVLRVRPT 186  
 121 PYVDATKSLWNDPGIOECYDRREYOLSDSTKYLLDLDVADPAVLPQODVLRVRPT 180  
 QY TGIIEYFPDQSYIFRNVGOGSRERKWHCFEENTSIIMFLVALSEYQVLESDNEN 246  
 187 TGIIEYFPDQSYIFRNVGOGSRERKWHCFEENTSIIMFLVALSEYQVLESDNEN 240  
 Db TGIIEYFPDQSYIFRNVGOGSRERKWHCFEENTSIIMFLVALSEYQVLESDNEN 240  
 181 TGIIEYFPDQSYIFRNVGOGSRERKWHCFEENTSIIMFLVALSEYQVLESDNEN 240  
 QY RMEESKALFRTITTYTFWFOSSVILFLNKKDLLEEKIMSHLVDPPEYDGPORDQAAR 306  
 247 RMEESKALFRTITTYTFWFOSSVILFLNKKDLLEEKIMSHLVDPPEYDGPORDQAAR 300  
 Db RMEESKALFRTITTYTFWFOSSVILFLNKKDLLEEKIMSHLVDPPEYDGPORDQAAR 300  
 241 RMEESKALFRTITTYTFWFOSSVILFLNKKDLLEEKIMSHLVDPPEYDGPORDQAAR 300  
 QY 307 EFLTKMFVDLPDSDKTIYSHFTCAIDTENTIREVFAAVKDTIIQLNLKEYNLV 359  
 301 EFLTKMFVDLPDSDKTIYSHFTCAIDTENTIREVFAAVKDTIIQLNLKEYNLV 353  
 Db 301 EFLTKMFVDLPDSDKTIYSHFTCAIDTENTIREVFAAVKDTIIQLNLKEYNLV 353

## RESULT 2

GBO\_CANFA STANDARD; PRT; 353 AA.  
 ID GBO\_CANFA  
 AC 028294;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Guanine nucleotide-binding protein G(q), alpha subunit.  
 GN GNAQ.  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 OX NCBI\_TaxID=9615;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96433124; PubMed=8836152;  
 RA Johnson G.J., Leis L.A., Dunlop P.C.;  
 RT "Specificity of G alpha q and G alpha 11 gene expression in platelets  
 and erythrocytes. Expressions of cellular differentiation and species  
 differences.";  
 RT Biochem. J. 318:1023-1031(1996).  
 RL Biochem. J. 318:1023-1031(1996).  
 CC -1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE  
 INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE  
 SIGNALING SYSTEMS.  
 CC -1- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).  
 THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.  
 CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(O)).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: L76257; AAB05548.1; ALT\_INIT.  
 DR HSSP: P10824; IBOF.  
 DR InterPro: IPR001019; Gprotelin\_alpha.  
 DR Pfam: PF00503; G-alpha.1.  
 DR PRINTS: PR00318; GPROTEIN.  
 DR PRODOM: PD000281; Gprotelin\_alpha.1.  
 DR SMART: SM00275; G-alpha.1.  
 KW GTP-binding; Transducer; Multigene family; ADP-ribosylation;  
 FT LIPID 3 3 PALMITATE (BY SIMILARITY).  
 FT LIPID 40 47 PALMITATE (BY SIMILARITY).  
 FT NP\_BIND 199 203 GTP (BY SIMILARITY).  
 FT NP\_BIND 268 271 GTP (BY SIMILARITY).  
 FT MOD\_RES 177 177 ADP-RIBOSYL(1) (BY ACTION OF CTX)  
 (BY SIMILARITY)  
 SQ SEQUENCE 353 AA: 41467 MW: 41737.071 E7373E07B1E4904C CRC64;

Query Match 98.0%; Score 1822; DB 1; Length 353;  
 Best Local Similarity 99.4%; Pred. No. 9.9e-130;  
 Matches 351; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 MACCLSEAEKARRINDEIRHVRKRDARRELKLLLTGSGSKSTFKOKRIIHSG 66  
 1 MACCLSEAEKARRINDEIRHVRKRDARRELKLLLTGSGSKSTFKOKRIIHSG 60  
 Db 1 MACCLSEAEKARRINDEIRHVRKRDARRELKLLLTGSGSKSTFKOKRIIHSG 60  
 QY YSDEKRGFTKLYONFTMQAMIRAMDTLKIPYKEHNKHAQAOLREVDEKVSFEN 126  
 61 YSDEKRGFTKLYONFTMQAMIRAMDTLKIPYKEHNKHAQAOLREVDEKVSFEN 120  
 Db YSDEKRGFTKLYONFTMQAMIRAMDTLKIPYKEHNKHAQAOLREVDEKVSFEN 120  
 61 YSDEKRGFTKLYONFTMQAMIRAMDTLKIPYKEHNKHAQAOLREVDEKVSFEN 120  
 QY PYVDATKSLWNDPGIOECYDRREYOLSDSTKYLLDLDVADPAVLPQODVLRVRPT 186  
 127 PYVDATKSLWNDPGIOECYDRREYOLSDSTKYLLDLDVADPAVLPQODVLRVRPT 180  
 Db PYVDATKSLWNDPGIOECYDRREYOLSDSTKYLLDLDVADPAVLPQODVLRVRPT 180  
 121 PYVDATKSLWNDPGIOECYDRREYOLSDSTKYLLDLDVADPAVLPQODVLRVRPT 180  
 QY TGIIEYFPDQSYIFRNVGOGSRERKWHCFEENTSIIMFLVALSEYQVLESDNEN 246  
 187 TGIIEYFPDQSYIFRNVGOGSRERKWHCFEENTSIIMFLVALSEYQVLESDNEN 240  
 Db TGIIEYFPDQSYIFRNVGOGSRERKWHCFEENTSIIMFLVALSEYQVLESDNEN 240  
 181 TGIIEYFPDQSYIFRNVGOGSRERKWHCFEENTSIIMFLVALSEYQVLESDNEN 240  
 QY RMEESKALFRTITTYTFWFOSSVILFLNKKDLLEEKIMSHLVDPPEYDGPORDQAAR 306  
 247 RMEESKALFRTITTYTFWFOSSVILFLNKKDLLEEKIMSHLVDPPEYDGPORDQAAR 300  
 Db RMEESKALFRTITTYTFWFOSSVILFLNKKDLLEEKIMSHLVDPPEYDGPORDQAAR 300  
 241 RMEESKALFRTITTYTFWFOSSVILFLNKKDLLEEKIMSHLVDPPEYDGPORDQAAR 300  
 QY 307 EFLTKMFVDLPDSDKTIYSHFTCAIDTENTIREVFAAVKDTIIQLNLKEYNLV 359  
 301 EFLTKMFVDLPDSDKTIYSHFTCAIDTENTIREVFAAVKDTIIQLNLKEYNLV 353  
 Db 301 EFLTKMFVDLPDSDKTIYSHFTCAIDTENTIREVFAAVKDTIIQLNLKEYNLV 353

## RESULT 3

GBO\_RAT STANDARD; PRT; 353 AA.  
 ID GBO\_RAT  
 AC P82471;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Guanine nucleotide-binding protein G(q), alpha subunit.  
 GN GNAO.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE=Brain;  
 RA Strohmann R.;  
 RT "Rat G alpha q subunit.";  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE OF 74-229 FROM N.A.  
 RC TISSUE=Brain cortex;  
 RX MEDLINE=93081611; PubMed=1333286;  
 RA Leshch K.-P., Manji H.K.;  
 RT "Signal-transducing G proteins and antidepressant drugs: evidence for  
 modulation of alpha subunit gene expression in rat brain.";  
 RT -----



RL Biol. Psychiatry 32:549-579(1992).  
 RN [3]  
 RP SEQUENCE OF 238-331 FROM N.A.  
 RC STRAIN-Sprague-Dawley; TISSUE-Kidney cortex;  
 RA Thomas C.P.;  
 RT "GTP-binding protein expression in glomerular mesangial cells";  
 RL Submitted (JUN-1995) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE SIGNALING SYSTEMS.  
 CC -1- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).  
 CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.  
 CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(0)).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: AF234260; AAF59930.1; ALT\_INIT.  
 DR EMBL: U37294; AAB02848.1; -.  
 DR HSSP: P10824; IBOF.  
 DR InterPro: IPR001019; Gprotein\_alpha.  
 DR Pfam: PF00503; G-alpha; 1.  
 DR ProDom: PD000281; Gprotein\_alpha; 1.  
 DR SMART: SM00275; G-alpha; 1.  
 DR GTP-binding; Transducer; Multigene family; ADP-ribosylation;  
 KW Palmitate; Lipoprotein.  
 KW Lipid  
 FT Lipid 3 3 PALMITATE (BY SIMILARITY).  
 FT NP\_BIND 4 4 GTP (BY SIMILARITY).  
 FT NP\_BIND 199 203 GTP (BY SIMILARITY).  
 FT NP\_BIND 268 271 GTP (BY SIMILARITY).  
 FT MOD\_RES 177 177 ADP-RIBOSYL[1] (BY ACTION OF CTX) (BY SIMILARITY).  
 FT CONFLICT 85 85 V -> I (IN REF. 2).  
 FT SEQUENCE 353 AA; 41469 MW; BBAC211FDD47534 CRC64;  
 SQ  
 Query Match 97.7%; Score 1818; DB 1; Length 353;  
 Best Local Similarity 98.9%; Pred. No. 2e-129;  
 Matches 349; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 Db 7 MACCSEAKKARRINDEIERHVRDKRARELKLILGSGSKPTFKOMRIIHSGS 66  
 1 MACCSEAKKARRINDEIERHVRDKRARELKLILGSGSKPTFKOMRIIHSGS 60  
 QY 67 YSDEDKRGFTKLVYQNIETAMQAMTRAMDTLKIPYKEHNKHAQIVREVDKYSAPFN 126  
 Db 61 YSDEDKRGFTKLVYQNIETAMQAMTRAMDTLKIPYKEHNKHAQIVREVDKYSAPFN 120  
 QY 127 PYVDAIKSLMNDPGIOECDYRRRQSLSTKYTYLLNDLDRVADPAYLPQODVLRVPT 186  
 Db 121 PYVDAIKSLMNDPGIOECDYRRRQSLSTKYTYLLNDLDRVADPAYLPQODVLRVPT 180  
 QY 187 TGIIEYPPDLSVIRFMDVGVGORSERRKMHCFENVTIMFLVALSEVDQVLESNDN 246  
 Db 181 TGIIEYPPDLSVIRFMDVGVGORSERRKMHCFENVTIMFLVALSEVDQVLESNDN 240  
 QY 247 RMEESKALERTIITYPWFQNSSVILFLNKKDLLEEKIMVSHLVDFEYDGPQRDAQAR 306  
 Db 241 RMEESKALERTIITYPWFQNSSVILFLNKKDLLEEKIMVSHLVDFEYDGPQRDAQAR 300  
 QY 307 EFTLKFVDLNDSDSKITISHTCATDTENIRFVFAVADTILQLNLKRYNY 359  
 Db 301 EFTLKFVDLNDSDSKITISHTCATDTENIRFVFAVADTILQLNLKRYNY 353  
 RESULT 4  
 GBO\_HUMAN  
 ID GBO\_HUMAN STANDARD: PRT; 353 AA.

AC P50148; Q13462; Q92471; O15108;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Guanine nucleotide-binding protein G(q), alpha subunit.  
 GN GNAQ OR GAO.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euteria; Primates; Catarrhini; Hominiidae; Homo.  
 OC NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96423032; PubMed=8825633;  
 RA Dong O., Shenker A., Way J., Haddad B.R., Lin K., Hughes M.R., McBride W.O., Spiegel A.M., Battey J.;  
 RT "Molecular cloning of human G alpha q cDNA and chromosomal localization of the G alpha q gene (GNAQ) and a processed pseudogene."; Genomics 30:470-475(1995).  
 RL [2]  
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.  
 RX TISSUE=prostate; PubMed=8664309;  
 RA Chen B., Leverette R.D., Schwin D.A., Kwatra M.M.;  
 RT "Human G(alpha q): cDNA and tissue distribution."; Biochim. Biophys. Acta 1281:125-128(1996).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96433124; PubMed=8836152;  
 RA Johnson G.J., Lela L.A., Dunlop P.C.;  
 RT "Specificity of G alpha q and G alpha 11 gene expression in platelets and erythrocytes. Expressions of cellular differentiation and species differences."; Biochem. J. 318:1023-1031(1996).  
 RL [4]  
 RP SEQUENCE FROM N.A.  
 RA Bai X.H., Acharya R., Rivera C., Murtagh J.J.;  
 RT Submitted (JUN-1997) to the EMBL/Genbank/DBJ databases.  
 RN [5]  
 RP SEQUENCE OF 74-229 FROM N.A.  
 RC TISSUE=Brain cortex;  
 RX MEDLINE=93081611; PubMed=1333286;  
 RA Leshch K.-P., Manji H.K.;  
 RT "Signal-transducing G proteins and antidepressant drugs: evidence for modulation of alpha subunit gene expression in rat brain."; Biol. Psychiatry 32:549-579(1992).  
 RN [6]  
 RP SEQUENCE OF 238-331 FROM N.A.  
 RC TISSUE=Hematopoietic;  
 RX MEDLINE=96077138; PubMed=7492305;  
 RA Thomas C.P., Dunn M.J., Mattern R.;  
 RT "Ca2+ signalling in K562 human erythroleukaemia cells: effect of dimethyl sulphoxide and role of G-proteins in thrombin- and thromboxane A2-activated pathways."; Biochem. J. 312:151-158(1995).  
 RL -1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE SIGNALING SYSTEMS.  
 CC -1- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).  
 CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.  
 CC -1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN OVARY, PROSTATE, TESTIS AND COLON.  
 CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(0)).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: U40038; AAC50363.1; ALT\_INIT.

DR EMBL: U43083; AAB06875.1; ALT\_INIT.  
 DR EMBL: L76256; AAB39498.1; ALT\_INIT.  
 DR EMBL: AF011496; AAB64301.1; ALT\_INIT.  
 DR EMBL: L40629; AAA99950.1; .  
 DR HSSP: P10824; IBOF.  
 DR Genew: HGNC:4390; GNAO.  
 DR MIM: 600998;  
 DR InterPro: IPR001019; Gprotein\_alpha.  
 DR Pfam: PF00503; G-alpha.1.  
 DR PRINTS: PR00318; GPROTEIN.  
 DR ProDom: PD000281; Gprotein\_alpha.1.  
 DR SMART: SM00275; G-alpha.1.  
 DR GTP-binding; Transducer; Multigene family; ADP-ribosylation;  
 KW Palmitate; Lipoprotein.  
 FT Lipid 3  
 FT LIPID 3  
 FT NP\_BIND 40 47  
 FT NP\_BIND 199 203  
 FT NP\_BIND 268 271  
 FT MOD\_RES 177 177  
 FT CONFLICT 22 23  
 FT CONFLICT 86 86  
 FT CONFLICT 97 97  
 FT CONFLICT 165 165  
 FT CONFLICT 318 318  
 FT CONFLICT 331 331  
 FT CONFLICT 352 352  
 FT CONFLICT 353 AA; 41425 MM; E779EB07B1F4904C CRC64;  
 SQ SEQUENCE

Query Match 97.7%; Score 1817; DB 1; Length 353;  
 Best Local Similarity 99.2%; Pred. No. 2,3e-129;  
 Matches 350; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 MACCLSEBAKERRINDEIERHVRKDKARREKLKLLGTGSGSKSTFIKQRIHSG 66  
 DB 1 MACCLSEBAKERRINDEIERHVRKDKARREKLKLLGTGSGSKSTFIKQRIHSG 60  
 QY 67 YSDEKRGFTKLVYQNTAMQAMIRAMDTLKIPKYEHKNAHQVREVEKVSAREN 126  
 DB 61 YSDEKRGFTKLVYQNTAMQAMIRAMDTLKIPKYEHKNAHQVREVEKVSAREN 120  
 QY 127 PYVAIKSLMNDPGIOECYDRRREYQLSDSTKYYLNDLRADPAYLPTOODVLRVPT 186  
 DB 121 PYVAIKSLMNDPGIOECYDRRREYQLSDSTKYYLNDLRADPAYLPTOODVLRVPT 180  
 QY 187 TGIIEYFPDLSVIFRWVDVGQSRERKWIHCENVTSLMFVALSEYDQVLVESDNE 246  
 DB 181 TGIIEYFPDLSVIFRWVDVGQSRERKWIHCENVTSLMFVALSEYDQVLVESDNE 240  
 QY 247 RMESKALFRITITYPWFQNSVILFLKKDLLEKIMYSLVDYFPETDGPORDAQAAR 306  
 DB 241 RMESKALFRITITYPWFQNSVILFLKKDLLEKIMYSLVDYFPETDGPORDAQAAR 300  
 QY 307 EFILKMFVDLNDPSDKITISHTCATOTENIRFEAAVKDITLONKEKYLAV 359  
 DB 301 EFILKMFVDLNDPSDKITISHTCATOTENIRFEAAVKDITLONKEKYLAV 353

RESULT 5  
 GBO\_XENLA STANDARD: PRT: 353 AA.  
 ID GBO\_XENLA  
 AC P38410.  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Guanine nucleotide-binding protein G(q), alpha subunit.  
 GN GNAO.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Laderer K.J., Smith L.D.;  
 RL Submitted (XXX-1992) to the EMBL/Genbank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Oocyte;  
 RX MEDLINE=94298961; PubMed=8026589;  
 RA Shapira H., Way J., Lipinsky D., Oron Y., Battey J.F.;  
 RT "Neurexins B receptor, expressed in Xenopus laevis oocytes,  
 RT selectively couples to G alpha q and not G alpha 11.";  
 RL FEBS Lett. 348:89-92(1994).  
 RN [3]  
 RP ERRTUM.  
 RA Shapira H., Way J., Lipinsky D., Oron Y., Battey J.F.;  
 RL FEBS Lett. 349:318-318(1994).  
 CC -1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE  
 CC INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE  
 CC SIGNALING SYSTEMS.  
 CC -1- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).  
 CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.  
 CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(O)).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: L05540; AAA9730.1; ALT\_INIT.  
 DR EMBL: U10502; AAB52189.1; ALT\_INIT.  
 DR HSSP: P10824; IBOF.  
 DR InterPro: IPR001019; Gprotein\_alpha.  
 DR Pfam: PF00503; G-alpha.1.  
 DR PRINTS: PR00318; GPROTEIN.  
 DR ProDom: PD000281; Gprotein\_alpha.1.  
 DR SMART: SM00275; G-alpha.1.  
 DR GTP-binding; Transducer; Multigene family; ADP-ribosylation;  
 KW Palmitate; Lipoprotein.  
 FT Lipid 3  
 FT LIPID 3  
 FT NP\_BIND 40 47  
 FT NP\_BIND 199 203  
 FT NP\_BIND 268 271  
 FT MOD\_RES 177 177  
 FT CONFLICT 10 10  
 FT CONFLICT 79 79  
 FT CONFLICT 158 158  
 FT CONFLICT 164 164  
 FT CONFLICT 353 AA; 41541 MM; 2310142412184553 CRC64;  
 SQ SEQUENCE

Query Match 94.9%; Score 1766; DB 1; Length 353;  
 Best Local Similarity 96.0%; Pred. No. 1.6e-125;  
 Matches 339; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 7 MACCLSEBAKERRINDEIERHVRKDKARREKLKLLGTGSGSKSTFIKQRIHSG 66  
 DB 1 MACCLSEBAKERRINDEIERHVRKDKARREKLKLLGTGSGSKSTFIKQRIHSG 60  
 QY 67 YSDEKRGFTKLVYQNTAMQAMIRAMDTLKIPKYEHKNAHQVREVEKVSAREN 126  
 DB 61 YSDEKRGFTKLVYQNTAMQAMIRAMDTLKIPKYEHKNAHQVREVEKVSAREN 120  
 QY 127 PYVAIKSLMNDPGIOECYDRRREYQLSDSTKYYLNDLRADPAYLPTOODVLRVPT 186  
 DB 121 PYVAIKSLMNDPGIOECYDRRREYQLSDSTKYYLNDLRADPAYLPTOODVLRVPT 180  
 QY 187 TGIIEYFPDLSVIFRWVDVGQSRERKWIHCENVTSLMFVALSEYDQVLVESDNE 246  
 DB 181 TGIIEYFPDLSVIFRWVDVGQSRERKWIHCENVTSLMFVALSEYDQVLVESDNE 240

QY 247 RMESEKALPRTITTYPMFQNSVILFLNKKDLLEEKIMTSHLVDFPEYDGPORDAQR 306  
 |||||||  
 DB 241 RMESEKALPRTITTYPMFQNSVILFLNKKDLLEEKIMTSHLVDFPEYDGPORDAQR 300  
 |||||||  
 QY 307 EFLKMFVLDNPSDKIITSHFTCATDTEINIRFVFAVADTIIQLNKEKNLV 359  
 |||||||  
 DB 301 EFLKMFVLDNPSDKIITSHFTCATDTEINIRFVFAVADTIIQLNKEKNLV 353  
 |||||||

## RESULT 6

GB11\_MELGA STANDARD; PRT: 359 AA.  
 ID GB11\_MELGA  
 AC P45645;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Guanine nucleotide-binding protein, alpha-11 subunit.  
 GN GN11.  
 OS Melagris gallopavo (Common turkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Meleagrididae; Melagris.  
 OX NCBI\_TaxID=9103;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC TISSUE=Brain;  
 RX MEDLINE=93207527; PubMed=8457205;  
 RA Maurice D.H., Waldo G.L., Morris A.J., Nicholas R.A., Harden T.K.;  
 RT "Identification of G alpha 11 as the phospholipase C-activating G-  
 protein of turkey erythrocytes.";  
 RL Biochem. J. 290:765-770(1993).  
 CC -1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE  
 INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE  
 SIGNALING SYSTEMS. ACTS AS AN ACTIVATOR OF PHOSPHOLIPASE C.  
 CC -1- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).  
 CC -1- THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.  
 CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(O)).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC DR EMBL: X73072; CAA51530.1;  
 CC DR HSSP: P10824; 1BOF.  
 CC DR Interpro: IPR001019; Gprotein\_alpha.  
 CC DR Pfam: PF00503; G-alpha\_1.  
 CC DR PRINTS: PR00318; GPROTEINA.  
 CC DR Prodom: PD00281; Gprotein\_alpha\_1.  
 CC DR SMART: SM00275; G-alpha\_1.  
 CC KW GTP-binding; Transducer; Multigene family; ADP-ribosylation.  
 CC FT NP\_BIND 46 53 GTP (BY SIMILARITY).  
 CC FT NP\_BIND 205 209 GTP (BY SIMILARITY).  
 CC FT NP\_BIND 274 277 GTP (BY SIMILARITY).  
 CC FT MOD\_RES 183 183 ADP-RIBOSYL[1] (BY ACTION OF CTX)  
 CC FT (BY SIMILARITY).  
 CC SQ SEQUENCE 359 AA: 42058 MW: B9E6427169BEAE CRC64:

Query Match 91.9%; Score 1709; DB 1; Length 359;  
 Best Local Similarity 90.3%; Pred. No. 3, 1e-121;  
 Matches 324; Conservative 22; Mismatches 13; Indels 0; Gaps 0;

QY 1 MLEESIMACCSSEKAEKARINDTERHNRKRRARLKLILGTSGSGSTPKQMR 60  
 |||||||  
 DB 1 MLEESIMACCSSEKAEKARINDTERHNRKRRARLKLILGTSGSGSTPKQMR 60  
 |||||||  
 QY 61 IHGSGYSDDEKRGFTKLYVONIFTAQAMIRAMPDLKIPYEHNRKAHQVREYDVK 120  
 |||||||  
 DB 61 IHGSGYSDDEKRGFTKLYVONIFTAQAMIRAMPDLKIPYEHNRKAHQVREYDVK 120  
 |||||||

QY 121 VSAFENPYDAIKSLMNDGICQYDRRREYOLSDSTKYLLNDLRVADPAVLPQOQVL 180  
 | | | | |  
 DB 121 VWFEPQPYSAKTLTMDNGIOECYDRREYOLSDSASAKYLSVDVRIATPGILPQOQVL 180  
 |||||||  
 QY 181 RVRVPTTGIIIEYFPDLQSVIFRMVDVGGORSERRKWHICFENVTSIMPLVALSEYDOVLV 240  
 |||||||  
 DB 181 RVRVPTTGIIIEYFPDLQSVIFRMVDVGGORSERRKWHICFENVTSIMPLVALSEYDOVLV 240  
 |||||||  
 QY 241 ESDNENRMESEKALPRTITTYPMFQNSVILFLNKKDLLEEKIMTSHLVDFPEYDGPQR 300  
 |||||||  
 DB 241 ESDNENRMESEKALPRTITTYPMFQNSVILFLNKKDLLEEKIMTSHLVDFPEYDGPQR 300  
 |||||||  
 QY 301 DAQAREFLKMFVLDNPSDKIITSHFTCATDTEINIRFVFAVADTIIQLNKEKNLV 359  
 |||||||  
 DB 301 DAQAREFLKMFVLDNPSDKIITSHFTCATDTEINIRFVFAVADTIIQLNKEKNLV 359  
 |||||||

## RESULT 7

GB11\_BOVIN STANDARD; PRT: 359 AA.  
 ID GB11\_BOVIN  
 AC P38409;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Guanine nucleotide-binding protein, alpha-11 subunit (GL2).  
 GN GN11.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91286303; PubMed=1905731;  
 RA Nakamura F., Ogata K., Shiozaki K., Kameyama K., Ohara K., Haga T.,  
 RT Nukada T.;  
 RT "Identification of two novel GTP-binding protein alpha-subunits that  
 RT lack apparent ADP-ribosylation sites for pertussis toxin.";  
 RL J. Biol. Chem. 266:12676-12681(1991).  
 CC -1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE  
 INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE  
 SIGNALING SYSTEMS.  
 CC -1- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).  
 CC -1- THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.  
 CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(O)).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC DR EMBL: D90336; BAA14350.1; ALT\_INIT.  
 CC DR PIR: B40891; B40891.  
 CC DR HSSP: P10824; 1BOF.  
 CC DR Interpro: IPR001019; Gprotein\_alpha.  
 CC DR Pfam: PF00503; G-alpha\_1.  
 CC DR PRINTS: PR00318; GPROTEINA.  
 CC DR Prodom: PD00281; Gprotein\_alpha\_1.  
 CC DR SMART: SM00275; G-alpha\_1.  
 CC KW GTP-binding; Transducer; Multigene family; ADP-ribosylation.  
 CC FT NP\_BIND 46 53 GTP (BY SIMILARITY).  
 CC FT NP\_BIND 205 209 GTP (BY SIMILARITY).  
 CC FT NP\_BIND 274 277 GTP (BY SIMILARITY).  
 CC FT MOD\_RES 183 183 ADP-RIBOSYL[1] (BY ACTION OF CTX)  
 CC FT (BY SIMILARITY).  
 CC SQ SEQUENCE 359 AA: 42070 MW: BOBAC4FBF5AAB8D5 CRC64:

Query Match 90.9%; Score 1691; DB 1; Length 359;  
 Best Local Similarity 89.4%; Pred. No. 6, 9e-120;  
 Matches 321; Conservative 24; Mismatches 14; Indels 0; Gaps 0;

Qy	1	MTTESIMACCLSEAKKARRINDEINEHNRDRDRARREIKLLLTGSGSGSTFIKQRR	60
Dp	1	MTLESMAACCLSEVKESKRINAEIKQRLRRDKRARREIKLLLTGSGSGSTFIKQRR	60
Qy	61	IIHSGSDDDKSGFTLVYQNIFTMOAMIRAMPDLKIPYKEHHKAAQALVREVDYER	12
Dp	61	IIHAGSISEEKRFETKLVYQNIFTAOAMIRAMELKLIYKEQKNAALLIREVDYER	12
Qy	121	VSAFENPYDAIKSLMNDPGIOECYDRRREYQJSDSTKYLYNDLDKVDVAPAYLPTQOQVL	18
Dp	121	VTFEEHRYVSAIKTLMNDPGIOECYDRRREYQJSDSAKYLYLNDVDAITSGLPTQOQVL	18
Qy	181	RVRPPTGIIEFPDQSVIERWADVGGQSEERKKIHCFENVTSMELVALSEYDOVLY	24
Dp	181	RVRPPTGIIEFPDLENIFRWADVGGQSEERKKIHCFENVTSMELVALSEYDOVLY	24
Qy	241	ESDNEENMEESKALFRTIITYPMFQNSVYLFTNKKDLLEEKIMVSHLVDFPEYDGPOR	30
Dp	241	ESDNEENMEESKALFRTIITYYPMFQNSVYLFTNKKDLLEDKILSHLVDFPEYDGPOR	30
Qy	301	DAQARREFILKMEVDLPPDSDKTIYSHFICATPTENIRVEFAAVKPTIILQNLKKEVNY	359
Dp	301	DAQARREFILKMEVDLPPDSDKTIYSHFICATPTENIRVEFAAVKPTIILQNLKKEVNY	359

RESULT 8	ID	GB11_HUMAN	STANDARD:	PRT:	359 AA.
AC	P29992	Q14350; O15109;			
DT	01-APR-1993	(Rel. 25, Created)			
DT	15-JUL-1998	(Rel. 36, Last sequence update)			
DT	16-OCT-2001	(Rel. 40, Last annotation update)			
DE	Guanine nucleotide-binding protein G(Y), alpha subunit (Alpha-11).				
GN	GNAL1 OR GAI1.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID:9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE:Retina;				
RX	MEDLINE:91219481; PubMed:1902575;				
RT	Jiang M., Pandey S., Tran V.T., Fong H.K.W.;				
RT	"Guanine nucleotide-binding regulatory proteins in retinal pigment				
RL	epithelial cells.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 88:3907-3911(1991).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RA	Bai X.H., Acharya R., Bai Y.H., Murtagh J.J.;				
RL	submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RA	Lamerlin J.E., McCreedy P.M., Skowronski E., Adamson A.W.,				
RA	Burkhardt-Schultz K., Gordon L., Kyle A., Ramirez M., Stillwagen S.,				
RA	Pian H., Velasco N., Do L., Regala W., Terry A., Garnes J.,				
RA	Dangaran L., Poundstone P., Christensen M., Georgescu A., Avila J.,				
RA	Liu S., Attix C., Andreise T., Trankhelm M., Amico-Keller G.,				
RA	Coeffield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G.,				
RA	Kronmiller B., Arellano A., Montgomerie M., Ow D., Nolan M., Trong S.,				
RA	Kobayashi A., Olsen A.S., Carraro A.V.;				
RL	submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.				
RN	[4]				
RP	SEQUENCE OF 244-337 FROM N.A.				
RC	TISSUE:Hematopoietic;				
RX	MEDLINE:96077138; PubMed:7492305;				
RT	Thomas C.P., Dunn M.J., Mattera R.;				
RT	"Ca2+ signalling in K562 human erythroleukaemia cells: effect of				
RT	dimethyl sulfoxide and role of G-proteins in thrombin- and				
RT	thromboxane A2-activated pathways.";				
RL	Biochem. J. 312:151-158(1995).				
CC	!- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE				
CC	INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE				

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR	EMBL:	M69013:	AAAS8624.1;	-.
DR	EMBL:	AF011497:	AAB64303.1;	-.
DR	EMBL:	AC005262:	AAC26415.1;	-.
DR	EMBL:	L40630:	AAA99949.1;	-.
DR	PIR:	A39394:	RGHUGY.	
DR	HSSP:	P10824:	IBOF.	
DR	Genew:	HGNC:4379:	GNAI1.	
DR	MIM:	139313:	-.	
DR	InterPro:	IPRO01019:	GproteIn_alpha.	
DR	Pfam:	PF00503:	G-alpha:1.	
DR	PRINTS:	PD00318:	GPROTEIN.	
DR	ProDom:	PD000281:	GproteIn_alpha:1.	
DR	SMART:	SM00275:	G-alpha:1.	
KW	GTP-binding:	Tiransducer:	Multigene family: ADP-ribosylation.	
FM	NP_BIND	46	53	GTP (BY SIMILARITY).
FT	NP_BIND	205	209	GTP (BY SIMILARITY).
FT	NP_BIND	274	277	GTP (BY SIMILARITY).
FT	MOD_RES	183	183	ADP-RIBOSYL[1] (BY ACTION OF CTX) (BY SIMILARITY).
FT	CONFLICT	6	6	M->I (IN REF. 2).
FT	CONFLICT	266	266	N->H (IN REF. 4).
FT	CONFLICT	285	285	Y->H (IN REF. 4).
FT	CONFLICT	301	302	DA->EP (IN REF. 1).
FT	CONFLICT	310	310	L->P (IN REF. 2).
QO	SEQUENCE	359 AA:	42123 MW:	DD3J71/6589E66046 CRC64;

[illegible]

DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Guanine nucleotide-binding protein, alpha-11 subunit.  
 GN GN11.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
 CC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC T1SSIDE-Oocyte;  
 RX MEDLINE=94298961; PubMed=8026589;  
 RA Shapira H., Way J., Lipinsky D., Oron Y., Battey J.F.;  
 RT "Neurexin B receptor, expressed in Xenopus laevis oocytes,  
 RT selectively couples to G alpha q and not G alpha 11.";  
 RL FEBS Lett. 348:89-92(1994).  
 RN [2]  
 RP ERRATUM.  
 RA Shapira H., Way J., Lipinsky D., Oron Y., Battey J.F.;  
 RL FEBS Lett. 349:318-318(1994).  
 CC -1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE  
 CC INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE  
 CC SIGNALING SYSTEMS. ACTS AS AN ACTIVATOR OF PHOSPHOLIPASE C.  
 CC -1- SUBUNIT: G PROTEIN IS COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).  
 CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.  
 CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(O)).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U10494; AA52188.1; -;  
 DR HSSP: P10824; IBOF.  
 DR InterPro: IPR001019; Gprotein\_alpha.  
 DR Pfam: PF00503; G-alpha.1.  
 DR PRINTS: PR00318; GPROTEIN.  
 DR ProDom: PD000281; Gprotein\_alpha.1.  
 DR SMART: SM00275; G-alpha.1.  
 DR GTP-binding; Transducer; Multigene family; ADP-ribosylation.  
 KW NP\_BIND 46 53 GTP (BY SIMILARITY).  
 FT NP\_BIND 205 209 GTP (BY SIMILARITY).  
 FT NP\_BIND 274 277 GTP (BY SIMILARITY).  
 FT MOD\_RES 183 183 ADP-RIBOSYL[1] (BY ACTION OF CTX) (BY  
 FT SIMILARITY).  
 SQ SEQUENCE 359 AA; 4208 MM; DA6B376993FDD870 CRC64;  
 Query Match 90.68; Score 1685; DB 1; Length 359;  
 Best Local Similarity 88.98; Pred. No. 1.9e-119;  
 Matches 319; Conservative 25; Mismatches 15; Indels 0; Gaps 0;

DB 241 ESDNENMEESKALFRTITTYPMFONSSVILFLNKKDLLEKIMYSHLVDFPEYDGPQR 300  
 QY 301 DAQAREFLIKMPDLPDSDKITYSHFCTATDRENRFAVAKOTIOLNKEVLY 359  
 DB 301 DAATAREFLIKMPDLPDSDKITYSHFCTATDRENRFAVAKOTIOLNKEVLY 359  
 RESULT 10  
 GB11\_MOUSE STANDARD; PRT; 359 AA.  
 ID GB11\_MOUSE  
 AC P21276; Q61939;  
 DT 01-MAY-1991 (Rel. 18, Created)  
 DT 01-MAY-1991 (Rel. 18, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Guanine nucleotide-binding protein, alpha-11 subunit.  
 GN GN11 OR GNA-11.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91067657; PubMed=2123549;  
 RA Strathmann M., Simon M.I.;  
 RT "G protein diversity: a distinct class of alpha subunits is present  
 RT in vertebrates and invertebrates.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:9113-9117(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/SV;  
 RX MEDLINE=96435432; PubMed=8838318;  
 RA Davidson I., Barnard M., Gavrilova O., Sweet K., Wilkie T.M.;  
 RT "Gene structure of murine Gna11 and Gna15: tandemly duplicated Gq  
 RT class G protein alpha subunit genes.";  
 RL Genomics 31:359-366(1996).  
 RN [3]  
 RP SEQUENCE OF 211-271 FROM N.A.  
 RX MEDLINE=90017488; PubMed=2508088;  
 RA Strathmann M., Wilkie T.M., Simon M.I.;  
 RT "Diversity of the G-protein family: sequences from five additional  
 RT alpha subunits in the mouse.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 86:7407-7409(1989).  
 CC -1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE  
 CC INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE  
 CC SIGNALING SYSTEMS. ACTS AS AN ACTIVATOR OF PHOSPHOLIPASE C.  
 CC -1- SUBUNIT: G PROTEIN IS COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).  
 CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.  
 CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(O)).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M55411; AAA63305.1; -;  
 DR EMBL: U37413; AAB36839.1; -;  
 DR EMBL: U37411; AAB36839.1; JOINED.  
 DR EMBL: U37412; AAB36839.1; JOINED.  
 DR EMBL: M57617; AAA63301.1; -;  
 DR PIR: B38414; RGM511.  
 DR PIR: B33833; B33833.  
 DR HSSP: P10824; IBOF.  
 DR MGD: MGI:95766; Gna11.  
 DR InterPro: IPR001019; Gprotein\_alpha.  
 DR Pfam: PF00503; G-alpha.1.  
 DR PRINTS: PR00318; GPROTEIN.  
 DR ProDom: PD000281; Gprotein\_alpha.1.  
 DR SMART: SM00275; G-alpha.1.  
 DR GTP-binding; Transducer; Multigene family; ADP-ribosylation.  
 KW NP\_BIND 46 53 GTP (BY SIMILARITY).  
 FT

```
FT NP_BIND 205 209 GTP (BY SIMILARITY).
FT NP_BIND 274 277 GTP (BY SIMILARITY).
FT MOD_RES 183 183 ADP-RIBOSYL(1) (BY ACTION OF CTX)
FT CONFLICT 77 77 K -> L (IN REF. 2).
SQ SEQUENCE 359 AA: 42024 MM: A33D2D6C6C6F85 CRC64:
Query Match 90.4%; Score 1682; DB 1; Length 359;
Best Local Similarity 88.6%; Pred. No. 3.3e-119;
Matches 318; Conservative 28; Mismatches 13; Indels 0; Gaps 0;

QY 1 MTLESTMACCLSEAKRARIINDEIERHVRDRDARREKLKLLTGSGSKSTFTQMR 60
DB 1 MTLESMAACCLSDVEESKRINAEIEKQLRDRKRDARREKLKLLTGSGSKSTFTQMR 60
QY 61 IIHGSGYSDDEKRGFTKLVQNFTAMQAMIRAMDTLKIPYKEHNKAHQAOLVREYDVK 120
DB 61 IIHGAGSEEDKRGFTKLVQNFTAMQAVRAMDTLKIKYKQNKANALLREVYDK 120
QY 121 VSAFENPYDAIKSLMNDPQIOECYDRRREYQSDSKYKYLNDLDRVADPAVLPQODVL 180
DB 121 VTTFEHQYVNAIKTLMSDPQVQECYDRRREFQSDSAKYLLTVDRTATVGYLPTQODVL 180
QY 181 RVAVPTGIIIEYFPDQSVIFRNVAVGQGSERRKWHCEENYTSIMFLVALSEYDQVL 240
DB 181 RVAVPTGIIIEYFPDLENIFRNVAVGQGSERRKWHCEENYTSIMFLVALSEYDQVL 240
QY 241 ESDNENMEESKALFRTIITYPWFQNSVILFLNKKDLLEKIMYSHLVDPPEYDQPR 300
DB 241 ESDNENMEESKALFRTIITYPWFQNSVILFLNKKDLLEKIMYSHLVDPPEYDQPR 300
QY 301 DAQAREFLIKMFVDLNPDSDKITTYSHFCATDPTENRFEFAVKDTIQLNKEYNLV 359
DB 301 DAQAREFLIKMFVDLNPDSDKITTYSHFCATDPTENRFEFAVKDTIQLNKEYNLV 359

RESULT 11
GB11_RAT STANDARD; PRT: 359 AA.
ID GB11_RAT
AC Q9JID2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Guanine nucleotide-binding protein, alpha-11 subunit.
GN GNA11.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Stroctman R.;
RT "Rattus norvegicus guanine nucleotide binding protein alpha 11 subunit
RT (G11).";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
CC INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
CC SIGNALING SYSTEMS.
CC -1- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(O)).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF239674; AAF81690.1;
CC DR HSSP; P10824; IBOF.
```

```
DR InterPro: IPR001019; Gprotein_alpha.
DR Pfam: PF00503; G-alpha_1.
DR PRINTS: PR00318; GPROTEINA.
DR ProDom: PD000281; Gprotein_alpha_1.
DR SMART: SM00275; G-alpha_1.
KM GTP-binding; Transducer; Multigene family; ADP-ribosylation.
FT NP_BIND 46 53 GTP (BY SIMILARITY).
FT NP_BIND 205 209 GTP (BY SIMILARITY).
FT NP_BIND 274 277 GTP (BY SIMILARITY).
FT MOD_RES 183 183 ADP-RIBOSYL(1) (BY ACTION OF CTX)
SQ SEQUENCE 359 AA: 42026 MM: B4CD057E9FC7092A CRC64:
Query Match 90.0%; Score 1674; DB 1; Length 359;
Best Local Similarity 88.0%; Pred. No. 1.3e-118;
Matches 316; Conservative 30; Mismatches 13; Indels 0; Gaps 0;

QY 1 MTLESTMACCLSEAKRARIINDEIERHVRDRDARREKLKLLTGSGSKSTFTQMR 60
DB 1 MTLESMAACCLSDVEESKRINAEIEKQLRDRKRDARREKLKLLTGSGSKSTFTQMR 60
QY 61 IIHGSGYSDDEKRGFTKLVQNFTAMQAMIRAMDTLKIPYKEHNKAHQAOLVREYDVK 120
DB 61 IIHGAGSEEDKRGFTKLVQNFTAMQAVRAMDTLKIKYKQNKANALLREVYDK 120
QY 121 VSAFENPYDAIKSLMNDPQIOECYDRRREYQSDSKYKYLNDLDRVADPAVLPQODVL 180
DB 121 VTTFEHQYVNAIKTLMSDPQVQECYDRRREFQSDSAKYLLTVDRTATVGYLPTQODVL 180
QY 181 RVAVPTGIIIEYFPDQSVIFRNVAVGQGSERRKWHCEENYTSIMFLVALSEYDQVL 240
DB 181 RVAVPTGIIIEYFPDLENIFRNVAVGQGSERRKWHCEENYTSIMFLVALSEYDQVL 240
QY 241 ESDNENMEESKALFRTIITYPWFQNSVILFLNKKDLLEKIMYSHLVDPPEYDQPR 300
DB 241 ESDNENMEESKALFRTIITYPWFQNSVILFLNKKDLLEKIMYSHLVDPPEYDQPR 300
QY 301 DAQAREFLIKMFVDLNPDSDKITTYSHFCATDPTENRFEFAVKDTIQLNKEYNLV 359
DB 301 DAQAREFLIKMFVDLNPDSDKITTYSHFCATDPTENRFEFAVKDTIQLNKEYNLV 359

RESULT 12
GB14_BOVIN STANDARD; PRT: 355 AA.
ID GB14_BOVIN
AC P38408;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DE Guanine nucleotide-binding protein, alpha-14 subunit (G11).
GN GNA14.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN
RP SEQUENCE FROM N.A.
RC MEDLINE=91286303; PubMed=1905731;
RA Nakamura F., Ogata K., Shiozaki K., Kameyama K., Ohara K., Haga T.,
RA Nakada T.;
RT "Identification of two novel GTP-binding protein alpha-subunits that
RT lack apparent ADP-ribosylation sites for pertussis toxin.";
RL J. Biol. Chem. 266:12676-12681(1991).
CC -1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
CC INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
CC SIGNALING SYSTEMS.
CC -1- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(O)).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; D90335; BAA1439.1; -  
DR PIR; A40891; A40891.  
DR HSSP; P10824; 1BOF.  
DR InterPro: IPR001019; G-protein\_alpha.  
DR Pfam; PF00503; G-alpha; 1.  
DR PRINTS; PR00318; GPROTEINA.  
DR ProDom; PD000281; G-protein\_alpha; 1.  
DR SMART; SM00275; G-alpha; 1.  
KW GTP-binding; Transducer; Multigene family; ADP-ribosylation.  
FT NP\_BIND 42 49 GTP (BY SIMILARITY).  
FT NP\_BIND 201 205 GTP (BY SIMILARITY).  
FT NP\_BIND 270 273 GTP (BY SIMILARITY).  
FT MOD\_RES 179 179 ADP-RIBOSYL[1] (BY ACTION OF CTX)  
FT (BY SIMILARITY).  
SQ SEQUENCE 355 AA; 41498 MW; 60886C6C9B7243B CRC64;  
  
Query Match 82.7%; Score 1538; DB 1; Length 355;  
Best Local Similarity 82.3%; Pred. No. 2.1e-108; Indels 0; Gaps 0;  
Matches 289; Conservative 33; Mismatches 29; Indels 0; Gaps 0;  
  
QY 9 CCLSEAEARINDEIRHVRDQKRDARREKLLLLGTGSGKSTFKOKRIIHSGSY 68  
DB 5 CCLSEAEKESQRISEIRQLRDKDKARREKLLLLGTGSGKSTFKOKRIIHSGSY 64  
QY 69 DEDKRGFTKLYONIFTFAMQAMIRAMDPLIKIPYKEHKNKAQOLREVDEKVSAPENPY 128  
DB 65 DEDKRGFTKLYONIFTFAMQAMIRAMDPLIKIQYCEQKKNQCLREVDEKVSAPENPY 124  
QY 129 VDAIKSLMNDPGIQCYRRREYQLSDSTKYYLNDLDRVADPAVLPDQDYLVRVPTTG 188  
DB 125 VEAIKQLMDEGIDCYRRREYQLSDSKAYLLDIDRIAMPAFPPDQDYLVRVPTTG 184  
QY 189 IIEYFEDQSYIFRVDVGGQSRERKMIHCFEENVTSMFLVALSEYDQVLAECNENRM 248  
DB 185 IIEYFEDLENIFRVDVGGQSRERKMIHCFESTSTIIFLVALSEYDQVLAECNENRM 244  
QY 249 EESKALFRTITTYPFQNSVYLFLNKKDLLEEKIMYSHLDYFPEYDGPORDAQAAREF 308  
DB 245 EESKALFRTITTYPFQNSVYLFLNKKDLLEEKIMYSHLDYFPEYDGPORDAQAAREF 304  
QY 309 ILKMFVDLNPDSKTIYSHFTCATDTENIRVFPAVKDTIQLNLEKNLV 359  
DB 305 ILKLYQDQNPDEKVIYSHFTCATDTENIRVFPAVKDTIQLNLEKNLV 355  
  
RESULT 13  
GB14\_XENLA STANDARD; PRT; 354 AA.  
ID GB14\_XENLA  
AC 073819;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Guanine nucleotide-binding protein, alpha-14 subunit.  
GN GNA14.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
OC Xenopodinae; Xenopus.  
OC NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98344029; PubMed=9677362;  
RA Shapira H., Amit I., Revach M., Oron Y., Batley J.F.;  
RT "Galphair 4 and Galphair mediate the response to trypsin in Xenopus  
oocytes".  
RL J. Biol. Chem. 273:19431-19436(1998).  
CC -!- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE

CC INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE  
CC SIGNALING SYSTEMS. ACTS AS AN ACTIVATOR OF PHOSPHOLIPASE C.  
CC MEDIATES RESPONSES TO TRYPSIN.  
CC -!- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).  
CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.  
CC -!- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 3 (G(0)).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AF059182; AAC1382.1; -  
DR HSSP; P10824; 1BOF.  
DR InterPro: IPR001019; G-protein\_alpha.  
DR Pfam; PF00503; G-alpha; 1.  
DR PRINTS; PR00318; GPROTEINA.  
DR ProDom; PD000281; G-protein\_alpha; 1.  
DR SMART; SM00275; G-alpha; 1.  
KW GTP-binding; Transducer; Multigene family; ADP-ribosylation.  
FT NP\_BIND 41 48 GTP (BY SIMILARITY).  
FT NP\_BIND 200 204 GTP (BY SIMILARITY).  
FT NP\_BIND 269 272 GTP (BY SIMILARITY).  
FT MOD\_RES 178 178 ADP-RIBOSYL[1] (BY ACTION OF CTX)  
FT (BY SIMILARITY).  
SQ SEQUENCE 354 AA; 41595 MW; C7105026B037600E CRC64;  
  
Query Match 82.6%; Score 1537; DB 1; Length 354;  
Best Local Similarity 80.9%; Pred. No. 2.5e-108; Indels 0; Gaps 0;  
Matches 284; Conservative 43; Mismatches 24; Indels 0; Gaps 0;  
  
QY 9 CCLSEAEARINDEIRHVRDQKRDARREKLLLLGTGSGKSTFKOKRIIHSGSY 68  
DB 4 CCLSEAEKESQRISEIRQLRDKDKARREKLLLLGTGSGKSTFKOKRIIHSGSY 63  
QY 69 DEDKRGFTKLYONIFTFAMQAMIRAMDPLIKIPYKEHKNKAQOLREVDEKVSAPENPY 128  
DB 64 DEDKRGFTKLYONIFTFAMQAMIRAMDPLIKIQTSTQNMENALVREVDEKVSLEKHH 123  
QY 129 VDAIKSLMNDPGIQCYRRREYQLSDSTKYYLNDLDRVADPAVLPDQDYLVRVPTTG 188  
DB 124 VEAIKQLMDEGIDCYRRREYQLSDSTKYYLNDLDRVADPAVLPDQDYLVRVPTTG 183  
QY 189 IIEYFEDQSYIFRVDVGGQSRERKMIHCFEENVTSMFLVALSEYDQVLAECNENRM 248  
DB 184 IIEYFEDLENIFRVDVGGQSRERKMIHCFEENVTSMFLVALSEYDQVLAECNENRM 243  
QY 249 EESKALFRTITTYPFQNSVYLFLNKKDLLEEKIMYSHLDYFPEYDGPORDAQAAREF 308  
DB 244 EESKALFRTITTYPFQNSVYLFLNKKDLLEEKIMYSHLDYFPEYDGPORDAQAAREF 303  
QY 309 ILKMFVDLNPDSKTIYSHFTCATDTENIRVFPAVKDTIQLNLEKNLV 359  
DB 304 ILKLYQDQNPDEKVIYSHFTCATDTENIRVFPAVKDTIQLNLEKNLV 354  
  
RESULT 14  
GB14\_MOUSE STANDARD; PRT; 355 AA.  
ID GB14\_MOUSE  
AC P30677;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Guanine nucleotide-binding protein, alpha-14 subunit.  
GN GNA14 OR GNA-14.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OC NCBI\_TaxID=10090;  
RN [1]





Db 121 QSVFAILSLMADAGIQECYDRRREYQLTDSAKYYLDAVDRIAEPNYLPTLQDILRVPT 180  
QY 187 TGIIEYFEDLOSIVIFRMVDYGGORSEERRKNIHCFENVTSMFLVALSEYDOVLVESDNEN 246  
Db 181 TGIIEYFEDLDSIFRMVDYGGORSEERRKNIHCFENVTSMFLVALSEYDOVLVESDNEN 240  
QY 247 RMESKALFRTIITYPWFONSSVLLFLNKKDLLEKIMYSHLVDPPEYDGPORDAQAAR 306  
Db 241 RMESKALFRTIITYPWFONSSVLLFLNKKDLLEKIMYSHLVDPPEYDGPORDAQAAR 300  
QY 307 EFILKMFVDLNPDSKTIYSHFTCATDTENIRFVFAAVKDTIQLNLKEYNLV 359  
Db 301 EFILKMFVDLNPDSKTIYSHFTCATDTENIRFVFAAVKDTIQLNLKEYNLV 353

Search completed: June 30, 2003, 16:16:29  
Job time : 13 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 30, 2003, 16:12:56 ; Search time 80 Seconds  
(without alignments)  
924.636 Million cell updates/sec

Title: US-09-899-295-2

Perfect score: 1860

Sequence: 1 MTELSIMACCLSEAKEARR.....VFNAVKDTILQLNLRKYNLV 359

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 segs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: SP archaea:\*  
2: SP bacteria:\*  
3: SP fungi:\*  
4: SP human:\*  
5: SP\_invertebrate:\*  
6: SP\_mammal:\*  
7: SP\_mhc:\*  
8: SP\_organelle:\*  
9: SP\_phage:\*  
10: SP\_plant:\*  
11: SP\_prodent:\*  
12: SP\_virus:\*  
13: SP\_vertebrate:\*  
14: SP\_unclassified:\*  
15: SP\_rvifus:\*  
16: SP\_bacteriap:\*  
17: SP\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
1	1849	99.4	359	4	09B2B9	09b2b9 homo sapien
2	1666	89.6	359	11	091X95	091x95 mus musculu
3	1544	83.0	353	5	P91955	P91955 limulus pol
4	1532	82.4	353	5	Q9U473	Q9u473 panulirus a
5	1524	81.9	355	11	08R2X9	08r2x9 mus musculu
6	1513	81.3	355	5	Q17386	Q17386 caenorhabdi
7	1504	80.9	353	5	Q9NL92	Q9nl92 octopus vul
8	1498	80.5	353	5	08T6P8	08t6p8 mamestra br
9	1439	77.4	360	5	08WSU8	08wsu8 halocynthia
10	1430	76.9	353	5	09NEZ0	09nez0 calliphora
11	1291	69.4	303	5	08T365	08t365 caenorhabdi
12	1152.5	62.0	305	5	Q9Y207	Q9y207 hydra magni
13	1122.5	60.3	355	5	Q9XZV4	Q9xzv4 geodia cydo
14	1112	59.8	279	5	09I7C8	09i7c8 drosophila
15	1055	56.7	374	6	09TUV3	09tuv3 oryctolagus
16	902	48.5	353	3	08TGA3	08tga3 leptosphaer

17	897.5	48.3	354	4	09UGA4	09uga4 homo sapien
18	897	48.2	353	3	09C134	09c134 tapesia yal
19	896	48.2	353	3	096VA7	096va7 fusarium ox
20	896	48.2	353	3	096T18	096t18 gibberella
21	895.5	48.1	354	13	090847	090847 gallus gall
22	893	48.0	353	3	09HFA2	09hfa2 rosellinia
23	892	48.0	353	3	096VW7	096vw7 trichoderma
24	887.5	47.7	354	13	090846	090846 gallus gall
25	885.5	47.6	354	4	08TAN5	08tan5 homo sapien
26	884.5	47.6	301	5	09Y202	09y202 ephydatia f
27	883.5	47.5	354	4	08TD72	08td72 homo sapien
28	883.5	47.5	354	11	09DC51	09dc51 mus musculu
29	882	47.4	305	5	Q9Y203	Q9y203 ephydatia f
30	881	47.4	469	5	08WP45	08wp45 halocynthia
31	875	47.0	355	13	09W6A4	09w6a4 squalus aca
32	874.5	47.0	354	3	09HEN1	09hen1 blumeria gr
33	874	47.0	353	3	09HEP9	09hep9 botrytis ci
34	873	46.9	350	13	09J365	09j365 xenopus lae
35	870.5	46.8	354	5	09NL93	09nl93 octopus vul
36	868	46.7	355	4	096C71	096c71 homo sapien
37	863.5	46.4	354	3	09HP99	09hfp9 schizophyil
38	861.5	46.3	354	13	013011	013011 xenopus lae
39	859	46.2	371	13	08OCY8	08ocy8 fuga rubrip
40	858.5	46.2	350	13	09Y123	09y123 ambystoma t
41	858.5	46.2	354	3	09C115	09c115 pisolithus
42	855.5	46.0	354	5	08WP45	08wp45 halocynthia
43	854.5	45.9	350	13	09OWX6	09owx6 brachydanio
44	854.5	45.9	354	13	09DC27	09dc27 gallus gall
45	851.5	45.8	354	5	08WSI1	08wsi1 clona inces

## ALIGNMENTS

RESULT 1  
ID 09B2B9 PRELIMINARY; PRT; 359 AA.  
AC 09B2B9;  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE G alpha q protein (Guanine nucleotide binding protein alpha q).  
GN GNAQ.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98366208; PubMed=9700850;  
RA Gabbeta J., Dhanasekaran N., Rao A.K.;  
RT "G alpha q cDNA sequence from human platelets.";  
RL Thromb. Res. 91:29-32(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Gabbeta J., Dhanasekaran N., Rao A.K.;  
RL submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Puri H.L. III, Ikeda S.R., Aronstam R.S.;  
RT "Homo sapiens guanine nucleotide binding protein (G protein) alpha q  
(GNAQ).";  
RL submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
EMBL: AF329284; AAC61117.1; -  
EMBL: AF493896; AAM12610.1; -  
HSSP: P10824; IBOF.  
DR InterPro: IPR001019; Gprotein\_alpha.  
DR Pfam: PF00503; G-alpha; 1.  
DR PRINTS: PR00318; GPROTEIN.  
DR PRODOM: PD000281; Gprotein\_alpha; 1.  
DR SMART: SM00275; G-alpha; 1.  
SO SEQUENCE 359 AA; 4214 MW; 6F69CAFF617DFA7C7 CRC64;

Query Match	Similarity	99.4%	Score 1849	DB 4	Length 359
Best Local	Similarity 99.4%	Pred. No. 36-136			
Matches 357	Conservative 1	Mismatches 1	Indels 0	Gaps 0	
QY	1	MTLESIMACCCSEEAKEARRINDEIERHVRDRDKRDARELKLILLGTGSESGKSTFIKQMR	60		
Db	1	MTLESIMACCCSEEAKEARRINDEIERLRDRKRDARELKLILLGTGSESGKSTFIKQMR	60		
QY	61	IIHSGSYDEDEKRGRTKLVYONIFPAMQAMIRAMDTLKIPYKEHNKAHQAOLVREVDK	120		
Db	61	IIHSGSYDEDEKRGRTKLVYONIFPAMQAMIRAMDTLKIPYKEHNKAHQAOLVREVDK	120		
QY	121	VSAFENPVPVDAIKSLMNDPGIOECYDRRREYQLSDSTYYLNDLDRVADPAVLPTQOVL	180		
Db	121	VSAFENPVPVDAIKSLMNDPGIOECYDRRREYQLSDSTYYLNDLDRVADPAVLPTQOVL	180		
QY	181	RVRVPTTGIIIEYPPDLOSIVIRAMVDVGQSRERRKWHICFENWTSIMFLVALSEYDOVL	240		
Db	181	RVRVPTTGIIIEYPPDLOSIVIRAMVDVGQSRERRKWHICFENWTSIMFLVALSEYDOVL	240		
QY	241	ESDNRNRKEESKALPRTITTYPMFONSSVILFLNKKDLLEKIMSHLVDFPEVDGQR	300		
Db	241	ESDNRNRKEESKALPRTITTYPMFONSSVILFLNKKDLLEKIMSHLVDFPEVDGQR	300		
QY	301	DAQAAREFILKMFVDLPNDPSDKIISYSHFTCATDTENIRFVPAVKDTILLQNLKEYNLV	359		
Db	301	DAQAAREFILKMFVDLPNDPSDKIISYSHFTCATDTENIRFVPAVKDTILLQNLKEYNLV	359		
RESULT 2					
Q91X85	PRELIMINARY:	PRT:	359	AA.	
ID	Q91X95				
AC	Q91X95:				
DT	01-DEC-2001 (TREMBLrel. 19, Created)				
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)				
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)				
DE	Similar to guanine nucleotide binding protein, alpha 11.				
GN	GNM1.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.				
OX	NCBI_Taxid:10090;				
RN	[1]				
RP	SEQUENCE FROM N. A.				
RC	TISSUE=COLON;				
RA	Strausberg R.;				
RL	Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; BC011169; AAH1169.1; -.				
DR	MGI; MGI:95766; Gna11.				
DR	InterPro; IPR001019; Gprotein_alpha.				
DR	Pfam; PF00503; G-alpha; 1.				
DR	Prodom; PD000281; Gprotein_alpha; 1.				
SO	SEQUENCE 359 AA; 41930 MW; 7437912F0BAC855 CRC64;				
Query Match					
	Best Local	Similarity 89.6%	Score 1666	DB 11	Length 359
	Matches 315	Conservative 29	Mismatches 15	Indels 0	Gaps 0
QY	1	MTLESIMACCCSEEAKEARRINDEIERHVRDRDKRDARELKLILLGTGSESGKSTFIKQMR	60		
Db	1	MTLESIMACCCSDYEKESKRINAELEKQLRDRKRDARELKLILLGTGSESGKSTFIKQMR	60		
QY	61	IIHSGSYDEDEKRGRTKLVYONIFPAMQAMIRAMDTLKIPYKEHNKAHQAOLVREVDK	120		
Db	61	IIHSGSYDEDEKRGRTKLVYONIFPAMQAMIRAMDTLKIPYKEHNKAHQAOLVREVDK	120		
QY	121	VSAFENPVPVDAIKSLMNDPGIOECYDRRREYQLSDSTYYLNDLDRVADPAVLPTQOVL	180		
Db	121	VTFTEHGVNVAIKTLMSDPGVOECYDRRREYQLSDSAKYLYTLVDRIATVGYLPTQOVL	180		
QY	181	RVRVPTTGIIIEYPPDLOSIVIRAMVDVGQSRERRKWHICFENWTSIMFLVALSEYDOVL	240		
Db	181	RVRVPTTGIIIEYPPDLOSIVIRAMVDVGQSRERRKWHICFENWTSIMFLVALSEYDOVL	240		
QY	240	VSFVPTTGIIIEYPPDLOSIVIRAMVDVGQSRERRKWHICFENWTSIMFLVALSEYDOVL	240		
Db	240	VSFVPTTGIIIEYPPDLOSIVIRAMVDVGQSRERRKWHICFENWTSIMFLVALSEYDOVL	240		
QY	240	VSFVPTTGIIIEYPPDLOSIVIRAMVDVGQSRERRKWHICFENWTSIMFLVALSEYDOVL	240		
Db	240	VSFVPTTGIIIEYPPDLOSIVIRAMVDVGQSRERRKWHICFENWTSIMFLVALSEYDOVL	240		

QY	241	ESNENRMEESKALPFTITITYMPFONSSVIFLNKKDLLEKIMYSLVYDPFPGQ	300
Db	241	ESNENRMEESKALPFTITITYMPFONSSVIFLNKKDLLEKIMYSLVYDPFPGQ	300
QY	301	DAQAAREFLIKMFVDLPDSDKIIYSHFTCATDTENIRFVFAAVKDTIQLNLKEYNLV	359
Db	301	DAQAAREFLIKMFVDLPDSDKIIYSHFTCATDTENIRFVFAAVKDTIQLNLKEYNLV	359
RESULT 3			
AC	P91955	PRELIMINARY;	PRT; 353 AA.
DT	01-MAY-1997 (T-EMBLrel. 03, Created)		
DT	01-MAY-1997 (T-EMBLrel. 03, Last sequence update)		
DT	01-MAY-2002 (T-EMBLrel. 20, Last annotation update)		
DE	Gq protein alpha subunit.		
OS	Limulus polyphemus (Atlantic horseshoe crab).		
OC	Eukaryotes; Metazoa; Arthropoda; Chelicerata; Merostomata; Xiphosura;		
OC	Limulidae; Limulus.		
OX	NCBI_TaxID=6850;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Ventral Eye.		
RA	Munger S.D., Schrammer-Berlin J.-L., Brink C.M., Battelle B.-A.;		
RT	"Molecular and Immunological Characterization of a Gq Protein from		
RT	Ventral and Lateral Eye of the Horseshoe Crab Limulus polyphemus.";		
RL	Invert. Neurosci. 0:0-0(1997).		
DR	EMBL; U08586; AAB48510.1; -.		
DR	HSSP; P10824; 1BOF.		
DR	InterPro; IPR001019; Gprotein_alpha.		
DR	Pfam; PF00503; G-alpha; 1.		
DR	PRINTS; PR00318; GPROTEINA.		
DR	ProDom; PD000281; Gprotein_alpha; 1.		
DR	SMART; SM00275; G-alpha; 1.		
SQ	SEQUENCE 353 AA; 41497 MW; C1C57783B3D2D516 CRC64; 4		
Query Match			
Best local Similarity 83.0%; Score 1544; DB 5; Length 353;			
Matches 293; Conservative 28; Mismatches 32; Indels 0; Gaps 0			
QY	7	MACCISEAKKARINDEIERHVRDKDARRELKLLLLGTGSGKSTFIKORRIIHSG	66
Db	1	MACCISEEEKKORINQEIETEROLRKDKDARELKLLLGTGSGKSTFIKORRIIHSG	60
QY	67	YSODEKRGCTKLVYONIFTAQAMIRADDTIKYKYEHNKAAHQAQREVDKVSAPEN	126
Db	61	YSDDKKSIIKILYONIIIMAAQMSNMKAMEMIKISKDRNNIENAEVLVSDYETVTFPS	120
QY	127	PYVDATKSLMNDPGIOECYDRRREYQSLDSKRYVLLNDLDRVADPAVLPDQDYLAVRP	186
Db	121	PVEAKTSLMVDPGIOECYDRRREYQSLDSKRYVLLNDLDRVADPAVLPDQDYLAVRP	180
QY	187	TGIIIEPPLQSVIRFMDVGGQSRERKKWIHCEENYTSIMFLVALSEYQVLVESDEN	246
Db	181	TGIIIEPPLDIIIRFMDVGGQSRERKKWIHCEENYTSIIFVALSEYQVLLVESDEN	240
QY	247	RMEESKALPFTITITYMPFONSSVIFLNKKDLLEKIMYSHLVYDPFPGQDPAQAR	306
Db	241	RMEESKALPFTITITYMPFONSSVIFLNKKDLLEKIMYSHLVYDPFPGQDPAQAR	300
QY	307	EFLIKMFVDLPDSDKIIYSHFTCATDTENIRFVFAAVKDTIQLNLKEYNLV	359
Db	301	EFLIKMFVDLPDSDKIIYSHFTCATDTENIRFVFAAVKDTIQLNLKEYNLV	353
RESULT 4			
ID	Q9U473	PRELIMINARY;	PRT; 353 AA.
AC	Q9U473;		
DT	01-MAY-2000 (T-EMBLrel. 13, Created)		
DT	01-MAY-2000 (T-EMBLrel. 13, Last sequence update)		

DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Gq/11 protein alpha subunit.  
 OS Panulirus argus (Spiny lobster).  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
 CC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Palinura;  
 OC Palinnoridea; Palinuridae; Panulirus.  
 OX NCBI\_TaxID=6737;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=OLFATORY ORGAN;  
 RX MEDLINE=20347127; PubMed=10781594;  
 RA Munger S.D., Gleeson R.A., Aldrich H.C., Rust N.C., Ache B.W.,  
 Greenberg R.M.;  
 RT "Characterization of a phosphoinositide-mediated odor transduction  
 pathway reveals plasma membrane localization of an inositol 1,4,5-  
 trisphosphate receptor in lobster olfactory receptor neurons."  
 RL J. Biol. Chem. 275:20450-20457(2000).  
 DR EMBL: AF201328; AAF19378.1; -.  
 DR HSSP: P10824; 1BOF.  
 DR InterPro: IPR001019; Gprotein\_alpha.  
 DR Pfam: PF00503; G-alpha.1.  
 DR PRINTS: PR00318; GPROTEINA.  
 DR Prodom: PD000281; Gprotein\_alpha.1.  
 DR SMART: SM00275; G-alpha.1.  
 SO SEQUENCE 353 AA; 41478 MW; FD9B551F66327BAA CRC64;

Query Match 82.4%; Score 1532; DB 5; Length 353;  
 Best Local Similarity 83.0%; Pred. No. 1,66-111;  
 Matches 293; Conservative 24; Mismatches 36; Indels 0; Gaps 0;

OY 7 MACCLSEBAKARRINDEIERHVRKRDARRELKLLLGTSKSTFIKOMRIHSGS 66  
 DB 1 MACCLSEBAKARRINDEIERHVRKRDARRELKLLLGTSKSTFIKOMRIHSGS 60  
 OY 67 YSDDKRGFTLYVONIFETAMQAMIRAMDITKIPYKYEHNKAHQVREVDVEXSAREN 126  
 DB 61 YSDDKRGFTLYVONIFETAMQAMIRAMDITKIPYKYEHNKAHQVREVDVEXSAREN 120  
 OY 127 PYVDAIKSLMNDPGIOECYDRRREYQSDSTKYVINDRVPALPTQODVLRVPT 186  
 DB 121 PYVDAIKSLMNDPGIOECYDRRREYQSDSTKYVINDRVPALPTQODVLRVPT 180  
 OY 187 TGIIEYPPDLQSVIFRWVDVGGGSRERKWIHCENVTSMFLVALSEYDOVLVESDNEN 246  
 DB 181 TGIIEYPPDLQSVIFRWVDVGGGSRERKWIHCENVTSMFLVALSEYDOVLVESDNEN 240  
 OY 247 RMESKALFRTITTYPMFQNSVILFLNKKDLLEKIMVSHLYVDFPEYDGPQRAQAR 306  
 DB 241 RMESKALFRTITTYPMFQNSVILFLNKKDLLEKIMVSHLYVDFPEYDGPQRAQAR 300  
 OY 307 EFLKMFVDLNPDSKTIYSHFTCATDTEENIRFVFAVAKDTIQLNKEYNLV 359  
 DB 301 EFLKMFVDLNPDSKTIYSHFTCATDTEENIRFVFAVAKDTIQLNKEYNLV 353

RESULT 5  
 OBR2X9 PRELIMINARY: PRT; 355 AA.  
 AC OBR2X9;  
 DT 01-JUN-2002 (TREMBLrel. 21, Created)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
 DE Guanine nucleotide binding protein, alpha 14.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Scuriognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC027015; AAH27015.1; -.  
 SO SEQUENCE 355 AA; 41527 MW; D34B39ACD179AE82 CRC64;

Query Match 81.9%; Score 1524; DB 11; Length 355;  
 Best Local Similarity 81.5%; Pred. No. 6,66-111;  
 Matches 286; Conservative 36; Mismatches 29; Indels 0; Gaps 0;

OY 9 CCCLSEBAKARRINDEIERHVRKRDARRELKLLLGTSKSTFIKOMRIHSGS 68  
 DB 5 CCCLSEBAKARRINDEIERHVRKRDARRELKLLLGTSKSTFIKOMRIHSGS 64  
 OY 69 DEDKRGFTLYVONIFETAMQAMIRAMDITKIPYKYEHNKAHQVREVDVEXSAREN 128  
 DB 65 DEDKRGFTLYVONIFETAMQAMIRAMDITKIPYKYEHNKAHQVREVDVEXSAREN 124  
 OY 129 VDAIKSLMNDPGIOECYDRRREYQSDSTKYVINDRVPALPTQODVLRVPT 188  
 DB 125 VDAIKSLMNDPGIOECYDRRREYQSDSTKYVINDRVPALPTQODVLRVPT 184  
 OY 189 TGIIEYPPDLQSVIFRWVDVGGGSRERKWIHCENVTSMFLVALSEYDOVLVESDNEN 248  
 DB 185 TGIIEYPPDLQSVIFRWVDVGGGSRERKWIHCENVTSMFLVALSEYDOVLVESDNEN 244  
 OY 249 ESKALFRTITTYPMFQNSVILFLNKKDLLEKIMVSHLYVDFPEYDGPQRAQAR 308  
 DB 245 ESKALFRTITTYPMFQNSVILFLNKKDLLEKIMVSHLYVDFPEYDGPQRAQAR 304  
 OY 309 ILKMFVDLNPDSKTIYSHFTCATDTEENIRFVFAVAKDTIQLNKEYNLV 359  
 DB 305 ILKMFVDLNPDSKTIYSHFTCATDTEENIRFVFAVAKDTIQLNKEYNLV 355

RESULT 6  
 O17386 PRELIMINARY: PRT; 355 AA.  
 AC O17386; 002546;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DE EGL-30 (HETEROTRIMERIC G protein alpha subunit).  
 GN EGL-30 OR M01D7.7.  
 OS Caenorhabditis elegans.  
 CC Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=N2;  
 RX MEDLINE=966221161; PubMed=8630258;  
 RA Brundage L., Avery L., Katz A., Kim U.J., Mendel J.E., Sternberg P.W.,  
 Simon M.I.;  
 RT "Mutations in a C. elegans Galpha gene disrupt movement, egg laying,  
 and viability."  
 RT Neuron 16:999-1009(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RX MEDLINE=94150718; PubMed=7906398;  
 RA Wilson R., Almscough R., Anderson K., Copsey T., Cooper J., Coulson A.,  
 Bonfield J., Burton J., Connell M., Copeley T., Cooper J., Coulson A.,  
 Craxton M., Dear S., Du Z., Durbin R., Pavello A., Fulton L.,  
 Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
 Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,  
 Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,  
 Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkhen R.,  
 Smalton N., Smith A., Sonhammer E., Steden R., Sulston J.,  
 Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
 Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 elegans."  
 RT Nature 368:32-38(1994).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RA Gattung S., Goela D., Wilson R.;



```
QY 127 PYVDAIKSLMNDPQIOECYDRRREYOLSDSTKYYLNDLDRVADPAVLPQOQVLRVRPT 186
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 121 PYVEAIKALWADAGIOCCYDRRREYQOLDSAKYLLQCIDRVAAPNVLPTQDILRVRVLT 180
QY 187 TGIIEPFDQSVIFRNVAVDGGORSEKRWIHCFENVTSTMFVALSEYDQVLESNEN 246
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 181 TGIIEPFDLEIRFRWADVGGORSEKRWIHCFENVTSTMFVALSEYDQVLESNEN 240
QY 247 RMESKALFRTITITYPFQNSVILFLNKKDLLEKIMYSHLVDPPEYDGPORDAQAAR 306
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 241 RMESKALFRTITITYPFQNSVILFLNKKDLLEKIMYSHLVDPPEYDGPORDAQAAR 300
QY 307 EFLIKMEVDLNPDSKTIYSHFTCATDTENIRFVFAAVKDTIQLNLEKYNL 358
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 301 EFLIRFVVDLNPDAEKIYSHFTCATDTENIKLVCAVKDTIMQSALKERNL 352
```

## RESULT 9

```
Q8WSU8 PRELIMINARY; PRT; 360 AA.
AC Q8WSU8;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE G-protein alpha subunit q class.
GN HRGQ.
OS Halocynthia roretzi (Sea squirt).
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OC Stolidobranchia; Pyuridae; Halocynthia.
OX NCBI_Taxid=7729;
RN [1]
RP SEQUENCE FROM N.A.
RA Iwasa T., Kanehara K., Watarai A., Ohkuma M., Tsuda M.;
RT "Cloning and expression pattern of G proteins in the larvae of
RT Halocynthia roretzi.";
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047084; BAB79199.1;
DR InterPro; IPR001019; Gprotein_alpha.
DR Pfam; PF00503; G-alpha.1.
DR PRINTS; PR00318; GPROTEIN.
DR ProDom; PD000281; Gprotein_alpha.1.
DR SMART; SM00275; G-alpha.1.
SQ SEQUENCE 360 AA; 42571 MW; A4ECFB6AE3C56DC9 CRC64;
```

Query Match 77.4%; Score 1439; DB 5; Length 360;

Best Local Similarity 78.3%; Pred. No. 2,9e-104;

Matches 282; Conservative 27; Mismatches 47; Indels 4; Gaps 2;

```
QY 3 LESTIMACLSSEAKERRINDEIERHVRDRKRDARRELKLLLGTSGSKSTFIKORRII 62
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 2 LEIMEFCCSEEOEAKRINAEIERELRHRRDRARRELKLLLGTSGSKSTFIKORRII 61
QY 63 HGSYSDEDKRGFKLYQNIPTAMQAMIRAMDTLIPYKEHNKAAQLVREVDKYS 122
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 62 HGSYSNDNRREFKLYQNIPTAMQAMIRAMDTLIPYKEHNKAAQLVREVDKYS 120
QY 123 AFEN---PYVDAIKSLMNDPQIOECYDRRREYOLSDSTKYYLNDLDRVADPAVLPQOQV 179
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 121 AEOGPFYCTALIEALMTDGIHDCYDRRREYQOLSDSTKYYLNDLDRVADPAVLPQOQV 180
QY 180 LVRVPTTGTIIEPFDQSVIFRNVAVDGGORSEKRWIHCFENVTSTMFVALSEYDQV 239
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 181 LVRVPTTGTIIEPFDQSVIFRNVAVDGGORSEKRWIHCFENVTSTMFVALSEYDQV 240
QY 240 VESNENRMESKALFRTITITYPFQNSVILFLNKKDLLEKIMYSHLVDPPEYDGPQ 299
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 241 VESNENRMESKALFRTITITYPFQNSVILFLNKKDLLEKIMYSHLVDPPEYDGPQ 300
QY 300 RDAQAAREFLIKMEVDLNPDSKTIYSHFTCATDTENIRFVFAAVKDTIQLNLEKYNL 359
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 301 QDANAAREFLIKMEVDLNPDSKTIYSHFTCATDTENIRFVFAAVKDTIQLNLEKYNL 360
```

## RESULT 10

```
Q9NEZ0 PRELIMINARY; PRT; 353 AA.
AC Q9NEZ0;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Guanine nucleotide-binding protein alpha subunit.
OS Calliphora vicina (Blue blowfly) (Calliphora erythrocephala).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Oestridae; Calliphoridae; Calliphora.
OX NCBI_Taxid=7373;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=EYE;
RA Schulz S., Huber A., Schwab K., Paulsen R.;
RT "A novel Ggamma isolated from Drosophila constitutes a visual G
RT protein gamma subunit of the fly compound eye.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ250443; CAB76453.1;
DR HSSP; P10824; IBOF.
DR InterPro; IPR001019; Gprotein_alpha.
DR Pfam; PF00503; G-alpha.1.
DR PRINTS; PR00318; GPROTEIN.
DR ProDom; PD000281; Gprotein_alpha.1.
DR SMART; SM00275; G-alpha.1.
SQ SEQUENCE 353 AA; 41295 MW; PFF799774C02072E CRC64;
```

Query Match 76.9%; Score 1430; DB 5; Length 353;

Best Local Similarity 76.1%; Pred. No. 1.4e-103;

Matches 268; Conservative 39; Mismatches 45; Indels 0; Gaps 0;

```
QY 7 MACCLSEAKERRINDEIERHVRDRKRDARRELKLLLGTSGSKSTFIKORRII 66
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1 MECLSEAKEQKRIINOETIERQRLRRDRARRELKLLLGTSGSKSTFIKORRII 60
QY 67 YSDEKRGFKLYQNIPTAMQAMIRAMDTLIPYKEHNKAAQLVREVDKYS AFEN 126
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 61 YSDDKRGFKLYQNIPTAMQAMIRAMDTLIPYKEHNKAAQLVREVDKYS AFEN 120
QY 127 PYVDAIKSLMNDPQIOECYDRRREYOLSDSTKYYLNDLDRVADPAVLPQOQVLRVRPT 186
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 121 PYLSAIRTLMADTQIOECYDRRREYOLDSAKYLLQCIDRVAAPNVLPTQDILRVRVLT 180
QY 187 TGIIEPFDQSVIFRNVAVDGGORSEKRWIHCFENVTSTMFVALSEYDQVLESNEN 246
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 181 TGIIEPFDQSVIFRNVAVDGGORSEKRWIHCFENVTSTMFVALSEYDQVLESNEN 240
QY 247 RMESKALFRTITITYPFQNSVILFLNKKDLLEKIMYSHLVDPPEYDGPORDAQAAR 306
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 241 RMESKALFRTITITYPFQNSVILFLNKKDLLEKIMYSHLVDPPEYDGPORDAQAAR 300
QY 307 EFLIKMEVDLNPDSKTIYSHFTCATDTENIRFVFAAVKDTIQLNLEKYNL 358
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 301 QFVLKXYLANPDEROCYSHFTTATDTENIKLVCAVKDTIMQNALKEFNL 352
RESULT 11
Q8T3G5 PRELIMINARY; PRT; 303 AA.
AC Q8T3G5;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE C. elegans Egl-30 protein (corresponding sequence M01D7.7b).
GN Egl-30.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
```





Qy	6	YSEDKRGKFKLYOYIFPAOMAMRAMDTLTKPKYEHKNAHMOVREYDVESAFEN	126
Db	61	YXQDCEKKNLVFRNIIIMSHMSMLQARTELKAIITDDAQRHQQMLALRPETASLGG	1200
Qy	127	PYVDAIKSLMNDPGIOECYDRREYQOLDSTKYUINDLDRVADPAVYLPFOQDVLVRPT	1866
Db	121	ETCEAEIRKIMDAGVOECYORRNEQOLSBSTKYUJLDDPLRITSSMDVPTQDVLVRPT	180
Qy	187	TGIIEXPRFDQSYIFPMVDVGQSRGRKMYICFENATSNIMFLVALSEYQDVLVESNE-	245
Db	181	TGIIIEPRFIINKIIRKMDVVGQSRGRKMYICFENATSNIMFLVALSEYQDVLVESNE-	240
Qy	246	NRMESKALFRITITYPWFENSSVILFLNKDLEEKIMYSHLVDPREYDQPDQAAA	305
Db	241	NRMVESLHLEFNTIISYPWENKSSILFLFNKKDLEEKYMHSHLIDYEEDYDGPCKDHVA	3000
Qy	306	REIIILMFADLNDPDSKIIITFHPGATQPTENIRPFAFVAKPTITQLNLKE	355
Db	301	REIIAMFETISINDMSADILPHTCATATPDKNEFEDVAKNHLIDQONITE	350

DT 01-MAR-2001 (TREMblrel. 16, Created)  
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE G-Protein alpha 49B  
 GN G-ALPHA-49B OR CG17759  
 OS *Drosophila melanogaster* (Fruit fly)  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha  
 OC Ephydroidea; Drosophilidae; *Drosophila*.  
 OX NCBI\_Taxid=7227;  
 GN [1]

RX MEDLINE=20196006; PubMed=10731132;

RA Adams W., Celniker S.E., Holt K.A., Evans C.A., Goekey J.D.,  
RA Anagnostis P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Asburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Rendall M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,  
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA April J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Baller J.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Boutch J., Brokstein P., Brotlier P.,  
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Dayes P.,  
RA de Pallos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodet A., Gong F., Gortell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Hayvey D., Helman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeigbam C.,  
RA Jaitai M., Kalush F., Kaplen G.H., Ke Z., Kemison J.A., Ketchum K.A.,  
RA Klumel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Mishina N.V., Mobery C., Morris J., Mostrel A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Murray D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclty J.M.,  
RA Palazzolo M., Peltan G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Styriass R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-T., Wassarman D.A., Weinstein G.M., Weissendach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye Y., Yen R.-F., Zaveril J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

Query Match	59.8%	Score 112:	DB 5:	Length 279;
Best Local Similarity	76.3%	Pred. No. 6.5e-79;		
Matches 213; Conservative	27;	Mismatches 19;	Indels 20;	Gaps 2

```

67 YSEBDKRGFKLVYQNIETAMQAMIRAMDILKIPY-KYEINKAHADLVREVDYEVSAFE 125
|||||: ||:|||| ||:||:|||| ||: ||: ||:

```

186 TTTGITEYPRFDLSVIFRMVDVGQGRSERRKNTHCFENWTSIMFLVALSEYDQVLVESDNE 245

QY	246 NRREESKALFRTIITYPMFONSSVILFLNKKDLEEKIM	284
	:         :   :        :	
Db	222 NRLEESKALFHTITTFEWFKNASILFLNKKMDVLEEKIM	260

DT 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
DE Galphal6.

05 Oryctolagus cuniculus (Rabbit).  
06 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
07 Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
08

NCBI\_TaxID=9986;

RP	RN	[1]	SEQUENCE FROM N.A.
----	----	-----	--------------------

RX MEDLINE=20035837; Pubmed=10571060;  
RA Feild J.A., Foley J.J., Testa T.T., Nuthallacanti P., Ellis C.,

RA Sarau H.M., Ames R.S.;  
 cloning and characterization of a rabbit ortholog of human Calpha16  
 PM

RT and mouse Gr1pha)15.8.

RL FEBS Lett. 460:53-56(1999).  
DR EMBL; AF169627; AAF06740.1; -.

DR HSP; P10824; 1AS3.  
DR InterPro: IPR001019: G-protein alpha.

DR Pfam; PF00503; G-alpha; 1.

DR PRINIS; PRO0316; GPROTEIN.A.  
DR ProDom; PD000281; Gprotein\_alpha; 1.

DR SMART; SM00275; G-alpha; 1.  
SEQUENCE 374 AA; 42933 MW; 630B3920A208CA1E CRC64;  
SQ

Query Match 56.7% Score 1055: DB 6: Length 374:

Best Local Similarity	55.88; Pred. No. 2.7e-74;
Match 203, Conservative	63, Mismatch 96, Indels 12, Gaps 3

Matches	% Conservative	% Misjudgments	% Indels	% Gaps
207	92	80	12	2

```
Db 13 CLSEDEKAAARVDQETITRLLLEHRRQYRGELKLLLGESGKSTFIKOMRIIHAGYSE 72
QY 70 EDRGFTKLYQNIIFTAMQAMIRAMDILKIPYKYEHNKAHAOLYREVDEKVSAPENPYV 129
Db 73 EDRGFRPLVQNIIFLSVQATIEAMDRLQIPYSRPEKLNASLVMSODPYKVNTEFETRYA 132
QY 130 DAISLNDPGIOECYDRRREYQUSDSTKYYLNDIDRVADPAILPTQODVLRVVPPTGI 189
Db 133 LAVQSLWRDAGVRACYEERREFFHLLDSAVYYLSHLERIAEBGYVPTAODVLRSMPTGI 192
QY 190 IEYFPDQSVIFRVDVGGQSRERKWIHCFENYTSIMFLVALSEYDQVLYESDNENRME 249
Db 193 NEYCFSVQKTNLRIVDVGQKSERKWIHCFEDVTALIFLASISEYDQCLEENGQENRMO 252
QY 250 ESKALFRITITYPWFONSSVILFLNKKDLLEKIMYSHLVDPPEYDQPOQDAQAREFI 309
Db 253 ESLALFCTVLAIPWFRATSVILFLNKTDILEDKVRTSHLATYFPGRGPPQDPEAKRFI 312
QY 310 LKMFVDL-----NPDG-----DKIYSHEFTCATDTENIRFVFAAVKDTLLQNLKEYN 357
Db 313 LELYTRYAGAAAGPDGASKGPRSRRLFSHYTCATDTONIRKVPKVDVDSVLARYLDEIN 372
QY 358 LV 359
Db 373 LL 374
```

Search completed: June 30, 2003, 16:17:58  
Job time : 83 secs